

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: October 28, 2003, 12:00:44

Search time 6.74545 Seconds  
(without alignments)  
2027.556 Million cell updates/sec

US-10-016-768A-1

1 KGRTPKXGKXNNDRLSLVE.....RAGSYGVPHSTLEKTKYKER 53

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SPREMBL 23:

- 1: sp archaea:
- 2: sp bacteria:
- 3: sp fungi:
- 4: sp human:
- 5: sp invertebrate:
- 6: sp mammal:
- 7: sp mhc:
- 8: sp organelle:
- 9: sp phage:
- 10: sp plant:
- 11: sp rodent:
- 12: sp virus:
- 13: sp vertebrate:
- 14: sp unclassified:
- 15: sp virus:
- 16: sp bacteriophage:
- 17: sp archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	278	100.0	1155	5 Q9V6D0	Q9V6D0 drosophila
2	275	98.9	1598	5 Q95YV8	Q95YV8 apis mellif
3	217	78.1	185	5 Q22051	Q22051 caenorhabdi
4	166	59.7	396	11 Q8C9Q0	Q8C9Q0 mus musculu
5	166	59.7	433	11 Q8BGT2	Q8BGT2 mus musculu
6	166	59.7	572	4 Q96JN0	Q96JN0 homo sapien
7	166	59.7	619	4 Q8N3I6	Q8N3I6 homo sapien
8	165	59.4	213	4 Q96NKL	Q96NKL mus musculu
9	165	59.4	517	11 Q8CUG4	Q8CUG4 mus musculu
10	99	35.6	1221	5 Q24079	Q24079 drosophila
11	92.5	33.3	645	5 Q8MKX3	Q8MKX3 drosophila
12	92.5	33.3	660	5 Q24457	Q24457 drosophila
13	92.5	33.3	1064	5 Q9V5N1	Q9V5N1 drosophila
14	92.5	33.3	1085	5 Q24455	Q24455 drosophila
15	84.5	30.4	661	5 Q9V8S2	Q9V8S2 drosophila
16	82	29.5	652	5 Q71168	Q71168 apis mellif

17	70.5	25.4	325	3 Q9VUG7	Q9VUG7 magnaporthe
18	70	24.3	158	17 Q26689	Q26689 methanobact
19	67.5	23.9	663	10 Q04976	Q04976 mangifera i
20	66.5	23.2	636	10 Q8LPL0	Q8LPL0 arabidopsis
21	64.5	23.2	728	10 Q9SCV0	Q9SCV0 arabidopsis
22	64.5	23.2	729	10 Q9S2I5	Q9S2I5 arabidopsis
23	64.5	23.2	737	10 Q8L5O9	Q8L5O9 citrus sine
24	64.5	23.2	532	3 Q92205	Q92205 botrytis ci
25	64	23.0	1046	5 Q9W0M2	Q9W0M2 drosophila
26	63.5	22.8	418	16 Q9B5H4	Q9B5H4 rhizobium i
27	63.5	22.8	721	10 Q9M5J4	Q9M5J4 phaseolus a
28	63.5	22.8	723	10 Q82670	Q82670 cicor ariet
29	63.5	22.7	368	17 Q97UG6	Q97UG6 sulfolobus
30	63	22.5	843	10 Q93X58	Q93X58 fragaria an
31	62.5	22.3	439	10 Q9SDK6	Q9SDK6 oryza sativ
32	62	22.3	324	12 Q41274	Q41274 spodoptera
33	61.5	22.1	378	10 Q04529	Q04529 arabidopsis
34	61.5	22.1	722	10 Q93X56	Q93X56 fragaria an
35	61.5	22.1	478	16 Q8R574	Q8R574 thermococ
36	61	21.9	739	5 Q8INS2	Q8INS2 drosophila
37	61	21.9	782	5 Q9V155	Q9V155 drosophila
38	61	21.9	948	2 Q8KQJ9	Q8KQJ9 saccharopol
39	61	21.9	528	2 Q9KVY9	Q9KVY9 buchnera ap
40	60.5	21.8	730	10 Q9ZPI7	Q9ZPI7 lupinus ang
41	60.5	21.8	731	10 Q9ZPI7	Q9ZPI7 pyrus pyrif
42	60.5	21.8	838	10 Q9ZPI1	Q9ZPI1 lycopersico
43	60.5	21.8	843	10 Q8L3P5	Q8L3P5 oryza sativ
44	60.5	21.6	100	2 Q9AFT0	Q9AFT0 shigella fl
45	60	21.6	100	2 Q9AFT0	Q9AFT0 shigella fl

## ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	1165 AA.
ID Q9V6D0	01-MAY-2000 (TREMBLrel. 13, Created)		
AC Q9V6D0	01-OCT-2002 (TREMBLrel. 22, Last sequence update)		
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE CG18389 protein.			
GN E1P93P OR CG18389.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophila.			
NCBI_TaxID=7227;			
OK NCBI_TaxID=7227;			
RN (1)			
RP SEQUENCE FROM N.A.			
RC STRAIN=Berkley;			
RX MEDLINE=20196006; PubMed=1071132;			
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA Amandas P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA Suton G.G., Mortan J.R., Yandell M.D., Zhang O., Chen L.X.,			
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe C.R., Miklos G.L.G.,			
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Baldwin D.,			
RA April J.F., Agdayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,			
RA Borkova D., Borkman M.R., Bouck J.P., Brokstein P., Brothier P.,			
RA Burtis K.C., Busam D.A., Butler H., Cadenot L.B., Davies P.,			
RA Cherry J.M., Cavley S., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA Durbin K.J., Evangelista C.C., Ferriz C., Ferris S., Fleischmann W.,			
RA Foster C., Gaborian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA Hareis N.L., Harvey D., Heiman T.J., Hernandez J.R., Honick J.,			
RA Hosten D., Houshun K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

**THIS PAGE BLANK (USPTO)**

Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
Mendelsohn G., Mielnik N.V., Mobarry C., Morris J., Mostoslavsky A.,  
Muller S.M., Moy M., Murphy B., Murphy L., Murzyn D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
Palazzo M., Pittman G.S., Sanders R.D.C., Scheeler F., Shen H.,  
Shen H., Spalding A.C., Stapleton M., Skupski M.P., Smith T.,  
Snyder B., Spalding A.C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang X., Weinstock G.M., Weissbach J., Wu D., Yang S., Yao Q.A.,  
Ye J., Yen R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
"The genome sequence of Drosophila melanogaster".  
Science 287:2185-2195(2000).

SEQUENCE FROM N.A.  
RA Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banazon J., An H., Baldwin D., Banazon J., Beeson K.Y., Busan D.A.,  
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
RA Dodson S., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fattian D.,  
RA Fectera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Iegwam C., Jalili M., Kruse D., Li P., Mallet B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Paclet J., Parag V., Park S., Patel S., Pfeiffer B., Scheeler F.,  
RA Phouamvong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Winkler S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RA "Sequencing of Drosophila melanogaster genome".  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.  
RA Mierra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniker S.E.,  
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Milburn G., Richter J., Ruoso S.,  
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RA "Annotation of Drosophila melanogaster genome".  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.  
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
RA "Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases".  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.  
RA Flybase;  
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RA EMBL: AL000373, AF55540.3, -  
RA Flybase: FB00013987, E1931F.  
RA GENE: 1165 AA; 123976 MW; A2556014070EBDBD CRC64;

Query Match  
Best Local Similarity 100.0%; Score 278; DB 5; Length 1165;  
Pred. No. 1.1e-24; Indels 0; Gaps 0;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
DB 758 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 810

RESULT 2  
08C97M8 PRELIMINARY; PRT: 1598 AA.  
AC 08C97M8;  
DT 01-MAR-2001 (TREMBlrel. 19, Created)  
DT 01-MAR-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 22, Last annotation update)

Mb1k-1 protein.  
GN MB1K-1.  
OS Apis mellifera (Honeybee).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis;  
OX NCBI\_TaxID=7460;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2187358; PubMed=11881813;  
RA Takuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohata M.,  
RA Fujikuni T., Kunita T., Sekimizu K., Natori S., Kubo T.,  
RT Identification of a novel gene, Mb1k-1, that encodes a putative  
RT transcription factor expressed preferentially in the large-type Kenyon  
RT cells of the honey bee brain".  
RL Insect Mol. Biol. 10:487-494(2001).  
DR EMBL: AB047034; BAB64310.1;  
SQ SEQUENCE 1598 AA; 174929 MW; ES475BD03ACBIEEF CRC64;

Query Match  
Best Local Similarity 98.1%; Score 275; DB 5; Length 1598;  
Pred. No. 3.6e-24;  
Matches 52; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
DB 1031 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 1083

RESULT 3  
ID 022051 PRELIMINARY; PRT: 185 AA.  
AC 022051;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE T01C1.3 Protein.  
CN T01C1.3  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloiderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Leonard N.,  
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RT none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
RT investigating biology".  
RT Science 282:2012-2018(1998).  
DR EMBL: Z68010; CAA92009.1;  
DR WormPep; T01C1.3; CE035594;  
SQ SEQUENCE 185 AA; 20706 MW; P9F59327B318F641 CRC64;

Query Match  
Best Local Similarity 78.1%; Score 217; DB 5; Length 185;  
Pred. No. 2.9e-18;  
Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

1 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
DB 83 KGTFRPGKGYRNDSDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 135

RESULT 4  
08C900 PRELIMINARY; PRT: 396 AA.  
AC 08C900;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Hypothetical protein (Fragment).

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 6.74545 Seconds  
(without alignments)  
2027.556 Million cell updates/sec

US-10-016-768A-1

Title: 278  
Percent score: 1 KGRPRKRGKRYNVDRLVE.....RAGSYGVPHSTLEKVKER 53  
Sequence: 1 KGRPRKRGKRYNVDRLVE.....RAGSYGVPHSTLEKVKER 53

Scoring table:  
Gapop 10.0 , Gapext 0.5

830525 segs, 258052604 residues

Number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP archaea:  
2: SP bacteria:  
3: SP fungi:  
4: SP human:  
5: SP invertebrate:  
6: SP mammal:  
7: SP mhc:  
8: SP organelle:  
9: SP plant:  
10: SP plant:  
11: SP rodent:  
12: SP virus:  
13: SP vertebrate:  
14: SP unclassified:  
15: SP virus:  
16: SP bacteriophage:  
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278	100.0	1165	5	Q9V6D0
2	275	98.9	1598	5	Q95YM8
3	217	78.1	185	5	Q22051
4	166	59.7	396	11	Q8C9Q0
5	166	59.7	433	11	Q8BGT2
6	166	59.7	572	4	Q96JNO
7	166	59.7	619	4	Q8N3L6
8	165	59.4	213	4	Q96NKL
9	165	59.4	517	11	Q8CUG4
10	99	35.6	1221	5	Q8MKX3
11	92.5	33.3	645	5	Q24457
12	92.5	33.3	660	5	Q24457
13	92.5	33.3	1064	5	Q9V5N1
14	92.5	33.3	1085	5	Q24455
15	84.5	30.4	661	5	Q9V8S2
16	82	29.5	652	5	Q77168

17	70.5	25.4	325	3	Q9VUG7	Q9VUG7 magnaporthe
18	70	25.2	393	11	Q8C9A6	Q8C9A6 mus musculus
19	67.5	24.3	158	17	Q26689	Q26689 methanobact
20	66.5	23.9	663	10	Q04976	Q04976 mangifera 1
21	64.5	23.2	636	10	Q8LPL0	Q8LPL0 arabidopsis
22	64.5	23.2	728	10	Q9SCV0	Q9SCV0 arabidopsis
23	64.5	23.2	729	10	Q9SZIS	Q9SZIS arabidopsis
24	64.5	23.2	737	10	Q8L509	Q8L509 citrus sine
25	64	23.0	532	3	Q92205	Q92205 botrytis ci
26	64	23.0	1046	5	Q9W0M2	Q9W0M2 drosophila
27	63.5	22.8	418	16	Q9W5J4	Q9W5J4 rhizobium 1
28	63.5	22.8	723	10	Q9M570	Q9M570 phaseolus a
29	63.5	22.8	723	10	Q9M570	Q9M570 cicer ariet
30	63	22.7	368	17	Q9JUG6	Q9JUG6 trifolobus
31	62.5	22.5	843	10	Q93X58	Q93X58 itegaria an
32	62	22.3	439	10	Q9SDK6	Q9SDK6 oryza sativ
33	61.5	22.1	324	12	Q41274	Q41274 spodoptera
34	61.5	22.1	378	10	Q04529	Q04529 arabidopsis
35	61.5	22.1	722	10	Q93X56	Q93X56 itegaria an
36	61	21.9	478	16	Q8R5T4	Q8R5T4 thermococ
37	61	21.9	739	5	Q8INS2	Q8INS2 drosophila
38	61	21.9	782	5	Q9V1S5	Q9V1S5 saccharopol
39	61	21.9	948	2	Q8KOL9	Q8KOL9 buchera ap
40	60.5	21.8	528	2	Q9KVV9	Q9KVV9 lupinus ang
41	60.5	21.8	730	10	Q9ZPI7	Q9ZPI7 pyrus pyril
42	60.5	21.8	731	10	Q9AYS1	Q9AYS1 lycopersico
43	60.5	21.8	838	10	Q9ZPI1	Q9ZPI1 oryza sativ
44	60.5	21.8	843	10	Q8L3P5	Q8L3P5 shigella fl
45	60	21.6	100	2	Q9AFT0	

## ALIGNMENTS

RESULT 1	ID	Q9V6D0	PRELIMINARY	PRT: 1165 AA.
AC	Q9V6D0	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	CG18389 Protein.			
GN	E1P93F OR CG18389.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RA	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton R.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,			
RA	Abdel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Bailly M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cusley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Fertile S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,			
RA	Gjoeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,			
RA	Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,			

**THIS PAGE BLANK (USPTO)**

RA. Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA. Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA. Mervulov G., Milstina N.V., Moberly C., Morris J., Moshirei A.,  
 RA. Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,  
 RA. Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paciel J.M.,  
 RA. Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA. Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA. Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA. Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA. Switkes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA. Wang Z.Y., Waasman D.A., Weinstein G.M., Weisenbach J.,  
 RA. Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA. Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA. Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA. Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).

[2]  
 RA. SEQUENCE FROM N.A.  
 RA. Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA. Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA. Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA. Carlson J.W., Center A., Chame M., Davenport L.B., Dietz S.M.,  
 RA. Dodson K., Doreet V., Doup L.E., Doyle C., Dresek D., Farfan D.,  
 RA. Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA. Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA. Idegawa C., Jalali M., Kruse D., Li P., Matzel B., Noshrefi A.,  
 RA. McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA. Paciel J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA. Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA. Stapleton M., Strong R., Switkes R., Tector C., Tyler D.,  
 RA. Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 "Sequencing of *Drosophila melanogaster* genome.";  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3]  
 RA. SEQUENCE FROM N.A.  
 RA. Mira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA. Hirdbeck P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA. Tupy J.L., Bergman C., Bernan B., Carlson J.W., Ceiniker S.E.,  
 RA. Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA. Krommiller B., Marshall B., Milburn B., Richter J., Russo S.,  
 RA. Seale S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA. Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 "Annotation of *Drosophila melanogaster* genome.";  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4]  
 RA. SEQUENCE FROM N.A.  
 RA. Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RA. SEQUENCE FROM N.A.  
 RA. Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR. EMBL: A8003737; AAF55940.3;  
 RA. Flybase; FBgn0013948; EId93F.  
 DR. Flybase; FBgn0013948; EId93F.  
 SQ. SEQUENCE 1165 AA; 123976 MW; A2556014070BEDD CRC64;

Query Match 100.0%; Score 278; DB 5; Length 1165;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-24;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGTTPKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 758 KGTTPKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 810

RESULT 2  
 ID 0895YM8 PRELIMINARY; PRT; 1598 AA.  
 AC 0895YM8;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Mb1k-1 protein.  
 GN MB1K-1.  
 OS Apis mellifera (Honeybee).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OC NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21873258; PubMed=11881813;  
 RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,  
 RA Fujiyuki T., Kumeda T., Sekimizu K., Natori S., Kubo T.;  
 RA "Identification of a novel gene, Mb1k-1, that encodes a putative  
 RA transcription factor expressed preferentially in the large-type Kenyon  
 RA cells of the honey bee brain.";  
 RT Insect Mol. Biol. 10:487-494(2001).  
 DR EMBL: AB047034; BAB64310.1;  
 SQ SEQUENCE 1598 AA; 174929 MW; E5475BDD3ACB1EEF CRC64;

Query Match 98.9%; Score 275; DB 5; Length 1598;  
 Best Local Similarity 98.1%; Pred. No. 3.6e-24;  
 Matches 52; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGTTPKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 1031 KGTTPKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 1083

RESULT 3  
 ID 022051 PRELIMINARY; PRT; 185 AA.  
 AC 022051;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE T01C1.3 protein.  
 GN T01C1.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lemard N.;  
 RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

[2]  
 RA. SEQUENCE FROM N.A.  
 RA. MEDLINE=99069613; PubMed=9851916;  
 RA. none;  
 RA. "Genome sequence of the nematode *C. elegans*: A platform for  
 RA. investigating biology.";  
 RT Science 282:2012-2018(1998).  
 DR. EMBL: Z68010; CAA92009.1;  
 DR. WormRep; T01C1.3; CE03594.  
 SQ. SEQUENCE 185 AA; 20706 MW; F9F593278318F641 CRC64;

Query Match 78.1%; Score 217; DB 5; Length 185;  
 Best Local Similarity 73.6%; Pred. No. 2.9e-18;  
 Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

OY 1 KGTTPKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 83 KSRPRKRGKRYNDPDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 135

RESULT 4  
 ID 08C9Q0 PRELIMINARY; PRT; 396 AA.  
 AC 08C9Q0;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).

**THIS PAGE BLANK (USPTO)**



PR 27-OCT-2000; 2000US-243865P.

XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.

XX Baehrecke EH;

XX WPI; 2002-479717/51.

PT Novel programmed cell death modulating proteins, useful for treating or  
PT preventing disorders associated with abnormal cell proliferation and  
PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
PT infarction

PS Claim 1; Fig 4; 88pp; English.

CC The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g., adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is human B93 programmed cell death modulating protein.

XX Sequence 442 AA;

Query Match 100.0%; Score 2250; DB 23; Length 442;

Best Local Similarity 100.0%; Pred. No. 3,2e-175; Mismatches 0; Indels 0; Gaps 0;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKMIROFAIEYISKSGTQENRNGSIQPSIVCKSIQNNQAEISLOEBOEGPLDITVNRM 60

DB 1 MKKMIROFAIEYISKSGTQENRNGSIQPSIVCKSIQNNQAEISLOEBOEGPLDITVNRM 60

QY 61 QEQNTQGGDGLVDLSTKTKTSIKSESSSICDPSSSESVAGRIHRNEDYVERSAEPADLL 120

DB 61 QEQNTQGGDGLVDLSTKTKTSIKSESSSICDPSSSESVAGRIHRNEDYVERSAEPADLL 120

QY 121 SKALKDIOSGALDINKAGILVIGIPKTKLLHLEALPAGKPSFKUKTRDPFDSYSYKDSK 180

DB 121 SKALKDIOSGALDINKAGILVIGIPKTKLLHLEALPAGKPSFKUKTRDPFDSYSYKDSK 180

QY 121 SKALKDIOSGALDINKAGILVIGIPKTKLLHLEALPAGKPSFKUKTRDPFDSYSYKDSK 180

DB 121 SKALKDIOSGALDINKAGILVIGIPKTKLLHLEALPAGKPSFKUKTRDPFDSYSYKDSK 180

QY 181 ETCAVLQKVALMARQAERTKSKINLLETSEIKFPTASTYLHOITLQKMTOPFEKNES 240

DB 181 ETCAVLQKVALMARQAERTKSKINLLETSEIKFPTASTYLHOITLQKMTOPFEKNES 240

QY 181 ETCAVLQKVALMARQAERTKSKINLLETSEIKFPTASTYLHOITLQKMTOPFEKNES 240

DB 181 ETCAVLQKVALMARQAERTKSKINLLETSEIKFPTASTYLHOITLQKMTOPFEKNES 240

QY 241 LOYTSNPTVOLKIPQLRVSSSVSKSPQDPGSGLDVWYVSVSKTSVLESGAOLKKNILPK 300

DB 241 LOYTSNPTVOLKIPQLRVSSSVSKSPQDPGSGLDVWYVSVSKTSVLESGAOLKKNILPK 300

QY 241 LOYTSNPTVOLKIPQLRVSSSVSKSPQDPGSGLDVWYVSVSKTSVLESGAOLKKNILPK 300

DB 241 LOYTSNPTVOLKIPQLRVSSSVSKSPQDPGSGLDVWYVSVSKTSVLESGAOLKKNILPK 300

QY 301 QNKIECSGPTVTHSSVDSYFLHGDLSPLCLNSKNGVDTSENTEGDLRKDSKOPRRKRG 360

DB 301 QNKIECSGPTVTHSSVDSYFLHGDLSPLCLNSKNGVDTSENTEGDLRKDSKOPRRKRG 360

QY 301 QNKIECSGPTVTHSSVDSYFLHGDLSPLCLNSKNGVDTSENTEGDLRKDSKOPRRKRG 360

DB 301 QNKIECSGPTVTHSSVDSYFLHGDLSPLCLNSKNGVDTSENTEGDLRKDSKOPRRKRG 360

QY 361 RYRGDHEIMEBAIWMVSGKMSVSKAAGIYGVPHSTLEKYKERSGTLKTPPKKRLP 420

DB 361 RYRGDHEIMEBAIWMVSGKMSVSKAAGIYGVPHSTLEKYKERSGTLKTPPKKRLP 420

QY 421 DTGLYNTDSTGSGCKNSKSPV 442

DB 421 DTGLYNTDSTGSGCKNSKSPV 442

RESULT 2

ID ABG17942 standard; Protein: 630 AA.

XX ABG17942;

AC 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17933.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS82129.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity

XX Claim 20; SEQ ID No 48301; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX and responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG30377 represent novel human

XX diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 630 AA;

Query Match 85.0%; Score 1913.5; DB 22; Length 630;

Best Local Similarity 84.1%; Pred. No. 1.6e-147; Mismatches 30; Indels 33; Gaps 4;

Matches 392; Conservative 11; Mismatches 30; Indels 33; Gaps 4;

QY 10 IEVYSKSGTQEN-----RNGSIGSIYCKSIQNNQAEISLOE-----48

DB 165 MELISQHDKVENKTIOTRRKROETLPAMRNSSDSMPFRQSLQIRGLASLDENTRK 224

QY 49 ---OEGLDITVNR-----MOEONTQGG---DGLVDLSTKTKTSIKSESSICDPSSENS 96

DB 225 KYTEKSSRKLTQNNNEISSDGEFYHQGNWQDVLVLSKTKTSIKSESSICDPSSENS 284

QY 97 VAGRLHRNEDYVERSAEAFADGLLSKALDIOSGALDINKAGILYGIPOKTLILLHLALP 156  
 DB 285 VAGRLHRNEDYVERSAEAFADGLLSKALDIOSGALDINKAGILYGIPOKTLILLHLALP 344  
 QY 157 ACKPASFKKKTDDPHDSYYSKSKETCANVLOKVALMARAOBERTESKINLLETSEIKRP 216  
 DB 345 ACKPASFKKKTDDPHDSYYSKSKETCANVLOKVALMARAOBERTESKINLLETSEIKRP 404  
 QY 217 TASTYLHQLTLOKMTQOFKKNESLQYETSNPTVOLKIPOLRVSYSKQOPGSGLLDM 276  
 DB 405 TASTYLHQLTLOKMTQOFKKNESLQYETSNPTVOLKIPOLRVSYSKQOPGSGLLDM 464  
 QY 277 YOVSKTSSVLEGSALQKLNILPKONKIECSGPVTHSSYDFLHDSPLCLNSKXGTV 336  
 DB 465 YOVSKTSSVLEGSALQKLNILPKONKIECSGPVTHSSYDFLHDSPLCLNSKXGTV 524  
 QY 337 DGTSENTEDEGLDRKSKOPRRKGRYROYDHEIMEEAIAMWSGKMSVSKAOGIYGVPHS 396  
 DB 525 DGTSENTEDEGLDRKSKOPRRKGRYROYDHEIMEEAIAMWSGKMSVSKAOGIYGVPHS 584  
 QY 397 TLEYKVKERSGTLTKPPKKRLPDTGLVNMMDSGTSGKSSKPV 442  
 DB 585 TLEYKVKERSGTLTKPPKKRLPDTGLVNMMDSGTSGKSSKPV 630

RESULT 3  
 ABP32451  
 ID ABP32451 standard; Protein: 104 AA.  
 AC ABP32451;  
 XX 09-JUL-2002 (first entry)  
 DT  
 XX Human ORF1424 protein, SEQ ID NO:2848.  
 DE  
 XX Human; ORF, open reading frame; ORFX; drug screening; diagnosis;  
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KW immune modulation; haematopoiesis regulation; tissue growth;  
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;  
 KW behaviour; cancer; proliferative disorder; neurological disorder;  
 KW cardiovascular disease; immune system disorder; organ transplantation;  
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 KW vasotrophic; antipariatic; antidiabetic; cytosatic; noctropic;  
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 KW cardiac; hypotensive; antihypoid; antiinflammatory; immunomodulator;  
 KW dermatological; analgesic; virucide; antibacterial; fungicide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200190366-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 24-MAY-2001; 2001WO-US17076.  
 XX  
 PR 24-MAY-2000; 2000US-206690P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Leach MD, Shimkets RA;  
 XX  
 DR WPI; 2002-106200/14.  
 XX  
 DR N-PSDB; ABN76477.  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and disorders related to organ  
 PT transplantation  
 PS Claim 10; Page 971-972; 2508pp; English.

XX Sequences ABP31028-ABP3561 represent 4334 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antiinfective activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.  
 CC  
 XX Sequence 104 AA:  
 SQ  
 QY Query Match 13.1%; Score 294.5; DB 23; Length 104;  
 DB Best Local Similarity 74.4%; Pred. No. 2.1e-16;  
 DB Matches 58; Conservative 7; Mismatches 8; Indels 5; Gaps 1;  
 QY 352 SKOPRRKGRYROYDHEIMEEAIAMWSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTK 411  
 DB 7 SKOPRRKGRYROYDHEIMEEAIAMWSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTK 66  
 QY 412 PPKKRLRL-----PDTGL 424  
 DB 67 PPKKRLRLMRSEGPVSV 84

RESULT 4  
 AAE24592  
 ID AAE24592 standard; Protein: 53 AA.  
 XX  
 AC AAE24592;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human E93 programmed cell death modulating protein conserved domain.  
 XX  
 KW Human; cancer; programmed cell death modulating protein; adenocarcinoma;  
 KW cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
 KW neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
 KW Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
 KW aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
 KW reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
 KW vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
 KW myeloma; noctropic; vasotrophic; immunostimulant; cerebroprotective;  
 KW cardiac; E93 protein.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200234882-A2.  
XX  
XX 02-MAY-2002.  
XX  
XX 29-OCT-2001; 2001WO-US48053.  
XX  
XX 27-OCT-2000; 2000US-243865P.  
XX  
XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
XX  
XX Baehrecke EH;  
XX  
XX WPI; 2002-479717/51.  
XX  
XX Novel programmed cell death modulating proteins, useful for treating or  
PT preventing disorders associated with abnormal cell proliferation and  
PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
PT infarction -  
XX  
XX Claim 1; Fig 1; 88pp; English.  
XX  
XX The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g. Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is human E93 programmed cell death modulating protein conserved  
CC domain.  
XX  
XX Sequence 53 AA;  
SQ  
Query Match 12.1%; Score 273; DB 23; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.6e-15;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 353 KQPKKKGRYQYDHEIMEEAIAVMWSGKMSVSKAGIGVPHSTLEYKVER 405  
DB 1 KQPKKKGRYQYDHEIMEEAIAVMWSGKMSVSKAGIGVPHSTLEYKVER 53  
RESULT 5  
AAE24593  
ID AAE24593 standard; Protein; 54 AA.  
XX  
XX AAE24593;  
AC  
XX  
XX 04-OCT-2002 (first entry)  
DT  
XX  
XX Fish E93 programmed cell death modulating protein conserved domain.  
DE  
XX  
XX Fish; cancer; programmed cell death modulating protein; adenocarcinoma;  
XX cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
XX neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
XX Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
XX aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
XX reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
XX vaccine; gene therapy; lymphoma; cytosstatic; melanoma; neuroprotective;  
XX myeloma; neutrotropic; vasotrophic; immunostimulant; cerebroprotective;  
XX cardiant; E93 protein.  
XX  
XX Tetraodon nigroviridis.

XX  
XX WO200234882-A2.  
XX  
XX 02-MAY-2002.  
XX  
XX 29-OCT-2001; 2001WO-US48053.  
XX  
XX 27-OCT-2000; 2000US-243865P.  
XX  
XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
XX  
XX Baehrecke EH;  
XX  
XX WPI; 2002-479717/51.  
XX  
XX Novel programmed cell death modulating proteins, useful for treating or  
PT preventing disorders associated with abnormal cell proliferation and  
PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
PT infarction -  
XX  
XX Claim 1; Fig 1; 88pp; English.  
XX  
XX The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g. Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is fish E93 programmed cell death modulating protein conserved  
CC domain.  
XX  
XX Sequence 54 AA;  
SQ  
Query Match 10.4%; Score 233.5; DB 23; Length 54;  
Best Local Similarity 81.5%; Pred. No. 8e-12;  
Matches 44; Conservative 7; Mismatches 2; Indels 1; Gaps 1;  
OY 353 KQPKKKGRYQYDHEIMEEA-IAVMWSGKMSVSKAGIGVPHSTLEYKVER 405  
DB 1 KQPKKKGRYQYDHEIMEEAITVMWSGKMSVSKAGIGVPHSTLEYKVER 54  
RESULT 6  
AAE24594  
ID AAE24594 standard; Protein; 53 AA.  
XX  
XX AAE24594;  
AC  
XX  
XX 04-OCT-2002 (first entry)  
DT  
XX  
XX Mouse E93 programmed cell death modulating protein conserved domain.  
DE  
XX  
XX Mouse; cancer; programmed cell death modulating protein; adenocarcinoma;  
XX cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
XX neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
XX Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
XX aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
XX reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
XX vaccine; gene therapy; lymphoma; cytosstatic; melanoma; neuroprotective;  
XX myeloma; neutrotropic; vasotrophic; immunostimulant; cerebroprotective;  
XX cardiant; E93 protein.  
XX

	Mus musculus.
PX	WO200234882-A2.
PD	02-MAY-2002.
PF	29-OCT-2001; 2001WO-US48053.
PR	27-OCT-2000; 2000US-24366SP.
PA	(UYWA-) UNIV MARYLAND BIOTECHNOLOGY INST.
PI	Baehtrecke EH;
DZ	WPI; 2002-479717/51.
PT	Noval programmed cell death modulating proteins, useful for treating or preventing disorders associated with abnormal cell proliferation and apoptosis such as cancer, stroke, Parkinson's disease, myocardial infarction -
PS	Claim 1; Fig 1; 86pp; English.
CC	The present invention relates to novel programmed cell death modulating proteins and polynucleotides encoding such proteins. Sequences of the invention are useful to screen potential cellular apoptosis inhibiting compounds to determine their use as therapeutic agents for treatment of diseases associated with increased programmed cell death. They are also useful for treating or preventing disorders associated with decrease in apoptosis. Programmed cell death modulating sequences are useful for treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma. Inhibition of the activity of the sequences of the invention are useful for treating disorders associated with increase in cell death or apoptosis such as acquired immunodeficiency syndrome (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic injuries (e.g., myocardial infarction, stroke, reperfusion injury), myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced diseases and other infectious or genetic immunodeficiencies. Sequences of the invention are used as vaccines and in gene therapy. The present sequence is mouse E93 programmed cell death modulating protein conserved domain.
SQ	Sequence 53 AA:
	Query Match            10.2%; Score 229; DB 23; Length 53; Best Local Similarity   81.1%; Pred.No. 1.8e+11; Matches      43; Conservative     6; Mismatches       4; Indels          0; Gaps                0.
OY	353 KOPKKRGRVRYOYDHEIMEEAIAMWSGKMSYSKAQGITYGVPHSTLEYKVKER 405  :: ::   ::::      Db        1 KHPPKRGRVRYNSEILEPISVLMSGKMSVKSKQSITYGIPHSLETKVNER 53
RESULT 7	
ABB7I145 ID	ABB7I145 standard; Protein; 1140 AA.
AC ABB7I145;	
DT DT	26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 40227.	
KW Drosophilid; developmental biology; cell signalling; insecticide; pharmaceutical.	
OS Drosophila melanogaster.	
FN WO200171042-A2.	
DD 27-SEP-2001.	

PF 23-MAR-2001; 2001WO-US09231.  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
DR WPI; 2001-656860/75.  
XX N-PSDB; ABL15248.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 40227; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABB57737-ABB72072).  
XX  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 1140 AA;

Query	March 1997	Score	200.5	DB	22	Length	1140			
	Best Local Similarity	21.7%	Pred. No.	3.1e-07						
	Matches	92	Conservative	66	Mismatches	146	Indels	119	Gaps	14
QY	19	TOENRNGSIGPSIVKSIQIMNOAENSLOEBOEGRLDILVNMNOBQNTQOQGDVLDL--ST	76							
DB	497	SGQNSNGSNASLLLDQQOQHQQHQQHQQOQQOQHVAAVRRRLPKSETPTNSSLDPNDAS	556							
QY	77	KKTSISESSSICDPSENSVAGRLHNRREDYVERSAEPADGLLSKALKDIOGALDINK	136							
DB	557	EDPILKIPFKVSGPSSSS-----LSP	579							
QY	137	AGLIVYIPKQTLILHLALPAKRPASPKTKTRDFHDSYKD-----SKETCAVLQKVALM	192							
DB	580	GGIVGG-----HHHPLNNNSLSISNNSN--HSSNHNGSNRRPHASPLAAV--	628							
QY	193	ARAQAERTKSKLNLLETSEIKFPASTYVLQTLTLOKVVQTFKKNESLOETSNPTVL	252							
DB	629	--AQGGYSAGNGLLVSSSSSIQKMAASIQOI-----NGSQGES-----	667							
QY	253	KIPQLRVSSVSKSGPDGSGGLDVMYQ-----VSTTSVLEGSALQIKLKNILPKQNKIEGS	307							
DB	668	---LNGNVSPCCSSNNGGSSSLGKPKPSISVAIKIGTPTSRPGASPNLLSOOH-----	718							
QY	308	GPVTHSSVDSYFLHGDLPFLCUNSKNGTVDGTSENTEEDGLDRKDS--KQPRKKRGARYOY	365							
DB	719	---HS--AHNL-----THQOQOQOQSLADEALCKGTPRPKGKRYNY	753							
QY	366	DHEIMEEALIAMVMSKMSVSKAOGIYGVPHSTLEYKVKERSGTLKTPPKKLLPPTGLY	425							
DB	754	DRDSLVEAVKAVQREGMSVHRAGSYGVPHSTLEYKVKENH---LMPRPKREKPPQDLY	810							
QY	426	NMT	428							
DB	811	GLT	813							

OY		137	AGILVGIPOKTLTLLHLEALPAKGPASFKNKTTRDFHDYSYKD-----SKETCAVLQKVALLW	192
Dd		597	GGLVVG-----HHHPLNNNNSLSINNSN--HSSNSHGNGSNRSPHSASPMLAAAV-	645
OY		193	ARQAERTESKUNLLETSEIKFPFPASTYLHQLTQLCKNVTFPKKNESLOYEISNPVQL	253
Dd		646	--AQGGYSAGNSLTITSSSSSIQKMASNIOROI-----NEQGQES-----	684
OY		253	KIPQLRVSSYSKSQPGDPSGLLDVMYQ-----VSKTSSVLEGSALOKKNLIPKONIECS	307
Dd		685	----LANGNVSDCCSNNNGSSSLGKRKRSISAKIIGTDISRGCASPILLSQH-----	735
OY		308	GPVTHSSVDSYFLPHGLPLCLNKKXGYVDGTSENTEDGLDKOS--KQPKRKGRRYOY	365
Dd		736	---HS---AHHL-----THQQOOOLSAOEALGKGTEPRKGRKYNY	770
OY		366	DHEIMEEATAMWSGKMSYSKAQGIYGVPHTSTLEKXYERSGTILKTPPKKLRLPDGLY	425
Dd		771	DRDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVERH--LMRPKRKEPPDPDLV	827
OY		426	NMT 428	
Dd		828	GLT 830	
<b>RESULT 9</b>				
ID	AAE24370	AAE24370 standard; protein; 53 AA.		
AC	AAE24370;			
XX				
Dt	04-OCT-2002	(first entry)		
DE				
XX	Fruit fly E93 programmed cell death modulating protein conserved domain.			
XX				
KM	Fruit fly; programmed cell death modulating protein; adenocarcinoma;			
KM	cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;			
KM	neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;			
KM	Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;			
KM	aplastic anemia; ischaemic injury; myocardial infarction; stroke;			
KM	reperfusion injury; toxin-induced disease; genetic immunodeficiency;			
KM	vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;			
KM	myeloma; nocitropic; vasotropic; immunostimulant; cerebroprotective;			
XX	cardiant; cancer; E93 protein.			
OS				
XX	Drosophila melanogaster.			
XX				
PN	WO200234882-A2.			
XX				
PD	02-MAY-2002.			
XX				
PE	29-OCT-2001; 2001MO-US48053.			
XX				
PR	27-OCT-2000; 2000US-243865P.			
XX				
PA	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.			
PI	Baehecke EH;			
DR	WPI; 2002-479717/51.			
PT				
PT	Novel programmed cell death modulating proteins, useful for treating or			
PT	preventing disorders associated with abnormal cell proliferation and			
PT	apoptosis such as cancer, stroke, Parkinson's disease, myocardial			
FT	infarction -			
XX				
PS	Claim 1; Fig 1; 88pp; English.			
XX				
CC	The present invention relates to novel programmed cell death modulating			
CC	proteins and polynucleotides encoding such proteins. Sequences of the			
CC	invention are useful to screen potential cellular apoptosis inhibiting			
CC	compounds to determine their use as therapeutic agents for treatment of			
CC	diseases associated with increased programmed cell death. They are also			



CC total A+T content that is less than about 70%. The BONT Ig protein is  
CC useful in vaccination against botulism, for eliciting protective immunity  
CC in a mammal, for treating dystonias, spasticity, pain, ocular motility,  
CC facial dyskinesias, stiff-person syndrome, bladder dysfunction, segmental  
CC myoclonus, hyperkinetic disorders, cosmetic treatment of facial wrinkles,  
CC conditions characterized by hyperactivity of the lower motor neuron, and  
CC to control autonomic nerve function or lipoe-walking due to stiff  
CC muscles common in children with cerebral palsy. The sequences are also  
CC useful for screening for botulinum neurotoxin inhibitors. This sequence  
CC represents a botulinum neurotoxin light chain serotype A protein.  
XX  
SQ Sequence 848 AA;  
  
Query Match 5.9%; Score 132.5; DB 23; Length 848;  
Best Local Similarity 20.4%; Pred. No. 0.074; Mismatches 129; Indels 115; Gaps 16;  
Matches 79; Conservative 64;  
  
QY 25 GSIGPSIVCKSIOMNQAENSLOEBOEGPLDITVNRMOBONTQOQDGVLDLSTKTSIKSE 84  
DB 273 GGHDPSPVISPSTDMNIVYKALQNPQD-----IANRLNIVSSAQSQGI-DILYKQIYKXK 326  
QY 85 ESSICDPSENSVAGRLHRNEDYERSAEFPADGLSLKADIOSGALDINKAGILYGP 144  
DB 327 YDFVEDPDKXYSV-----DKDKF-----DLKYALMFGFTETWLAG-EYGI- 366  
QY 145 OKTLLHL-EALP-----AGKPASFKKKTROFH-----DSYSYKDSKE 181  
DB 367 -KTRYSYSEYLPFKTEKLLDNTIYTONEGFNINASKULKTEFPNQKAVNKAEEISL 425  
QY 182 TCAVLQKVALMARAQERTSKSLNLTSEIKFPTASTYVLIHQVMTQFKEKNESL 241  
DB 426 EHLVIYIRAMCKPWPVKYTKGSEQCIIYNNEDLFFIAN-----KQSFKDLAKMTI 477  
QY 242 QYFNSNPVQ-----LKITP-QLRVSSYSKQP 267  
DB 478 AYNTQNTNIENNFSIDQLINDLSGGIDLPNENTPEPTNDDIDIPYIKQSAKIKFV 537  
QY 268 DSGGLDMYGVSKTSVLEGSALQKKNILPKXNK-----IEGSGVTSSVDSY 318  
DB 538 DGDSLFEYLAHQTPPSNI-ENQLTNSLNDALRNKKYTPFSTNLVERKANTVVGAS---- 592  
QY 319 FLHGDLSPLCLNSKNGTVDG-TSENTE 344  
DB 593 -----LFVNWVGVIDDFTSESTQ 611  
  
RESULT 12  
AAR9795 ID AAR9795 standard; Protein; 3248 AA.  
XX  
XX AAR9795;  
XX AC  
XX 08-OCT-1996 (first entry)  
DT 08-OCT-1996 (first entry)  
XX  
DE Kinetochore protein CENP-F.  
XX  
KM Kinetochore protein; CENP-F; cell cycle; cancer; diagnosis;  
KW autoimmune antibody.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Domain 1..200  
FT Domain /label= Extended\_coiled\_structure  
FT Domain 280..1350  
FT Domain /label= Extended\_coiled\_structure  
FT Domain 1380..1610  
FT Domain /label= Globular\_domain  
FT /note= "globular domain consists of 2 direct  
FT repeats of 95 amino acids"  
FT Domain 1620..1750  
FT /label= Extended\_coiled\_structure  
FT Domain 1850..2990

FT FT  
FT Domain /label= Extended\_coiled\_structure  
FT 3048..3248  
FT /label= C-terminal domain  
FT /note= "the C-terminal domain is predicted to  
FT form a proline-rich (10.6%) highly  
FT basic (pI 10) globular domain"  
XX  
XX MO9617867-A1.  
XX  
XX 13-JUN-1996.  
XX  
XX 08-DEC-1995; 95WO-US16216.  
XX  
XX 09-DEC-1994; 94US-0353700.  
XX  
XX (FOX-C-) FOX CHASE CANCER CENT.  
XX (UYTE-) UNIV TECHNOLOGIES INT INC.  
XX  
XX Rattner JB, Yen TJ;  
XX WPI, 1996-287116/29.  
XX DR N-PSDB; AAT34578.  
XX  
XX DNA encoding kinetochore protein - used as a marker for the G2 and M  
XX phases of a cell cycle, partic. for detection of malignant diseases  
XX  
XX Claim 12; Page 41-54; 72pp; English.  
XX  
XX A 372 kDa human kinetochore protein, CENP-F (AAR9795), is detected  
XX by immunofluorescence microscopy only during the G2 and M phases  
XX of a cell cycle. It is the product of a cDNA clone (AAT34578)  
XX isolated from a breast carcinoma cDNA library. Recombinant CENP-F  
XX can be produced by expression in prokaryotic or eukaryotic host  
XX cells. CENP-F can be used to detect autoimmune antibodies to  
XX the protein, which may provide an early diagnosis for the onset  
XX of various malignant diseases. Use of CENP-F as a cell cycle  
XX marker allows the specific detection of G2 and M phase cells.  
SQ Sequence 3248 AA;  
  
Query Match 5.8%; Score 130; DB 17; Length 3248;  
Best Local Similarity 20.2%; Pred. No. 0.8; Mismatches 200; Indels 138; Gaps 21;  
Matches 104; Conservative 74;  
  
QY 7 QFATFYTSKSGKTOENR---NGSIGPSIVCKSIOMNQAENSLOEBOEGPLDITVNRMOEQ 63  
DB 2676 QDTLEVLQSSYKXNLELTKMDKMSFVEKVNMTAKETELQREHMAQKTAELQBEL 2735  
QY 64 NTOQGDGVLDLSTKTSIKSEESICDPSENS-VAGRLHRNEDYERSAEFADGL--- 119  
DB 2736 SGEKNRLAGELQILLIEIKSSKQDLKELTLENSLKLSCOMHKDOYEKKGKVAEYAEY 2795  
QY 120 ---LSKALKDIQSGALDINKAGILYGIPOKTLHLHLBALPAGKPASFKNTRDPHDSYSY 176  
DB 2796 QLRHBAEKHQALLLDTNKK--YVEVIQT-----YREKL----- 2828  
QY 177 KDSKETCAVLQKVALMARAQERTSKLN--LLETSEIKFPTASTYVLIHQVMTQF 234  
DB 2829 -TSKEECLSSQKLEI---DLKSSKELNNSLKATQILIEELKTKTMDNL---KYVNL 2880  
QY 235 KENKESL-----QYETSNPVQKIPOLRVSSVSKSQPDGSLDVMYGVSKT 282  
DB 2881 KKENERAQGMKMLIKSCQLEKEKELQKELQSLQAAQ-----EKQKT 2924  
QY 283 SSYLE-----GSAIQKKNILPKXNKI-----CSGVTSSVD--SYFLHGLSPL 327  
DB 2925 GTVMDTVDELITBRIKELKTEKTEADDEVLDKYSLLISHEKLEKAKEMETGYAHL 2984  
QY 328 C-----LNSKNGTVDG-----TSENTEGDIDRDSKQPRKK-----RGYRQ 364  
DB 2985 CSQGSKQDSRSGPLGIVVPSPSPISVTEKRLSSGQNKASGKRQSSGIWENGRGPTPA 3044  
QY 365 YDHEIMEALAMWWSGKSVSKAQGI-----YGVPH-----STLEYKVKENSGT 408

```

Db      3045 TRESFSKSKKAWMSGIHAEDETEGEPEPEPEGLPEVVKKGAFADIPGTGKSPYLRTTWA 3104
Oy      409 LKTPPK---KKLRLPDPTGLYNNMTDSGTSCXNSKSP 441
Db      3105 TTSFRLAAQKALSLPLSL-----GKENVLAESSKP 3134

RESULT 13
AAU82977
ID      AAU82977 standard; Protein; 533 AA.
AC      AAU82977;
XX      23-APR-2002 (first entry)
XX      S. cerevisiae BFR2 protein target for antifungal compound.
XX      DE
XX      Anti-fungal; fungal gene transcription; RPC34; POP3; TFA2; NAB2;
XX      MW: MTR2; BOS1; POL30; RSA2; SQT1; MTW1; TFB1; SPC98; BFR2; RNAL;
XX      GCD1; SKI6; NIP1; LCP5; NCE103; ECO1; ORC2; CNS1; YPD1; TIM10; SRB4;
XX      yeast; fungus.
XX      OS      Saccharomyces cerevisiae.
XX      PN      WO200202055-A2.
XX      PD      10-JAN-2002.
XX      PF      28-JUN-2001; 2001WO-US20592.
XX      PR      29-JUN-2000; 2000US-215164P.
XX      PS      10-AUG-2000; 2000US-224457P.
XX      PA      (AAND-) ANADYS PHARM INC.
XX      PI      Moore J, Buurman ET, Desilva T, Harris S, Komarnitsky S;
XX      Mendillo M, Moore D, McCoy M, Sanderson K, Haq T, Zhu S, Long F;
XX      Davidov E, Thompson CM;
XX      DR      WPI; 2002-147962/19.
XX      N-PSDB; ABRK32865.
XX      PT      Screening candidate antifungal compound for interaction with essential
XX      PT      protein, modulation of essential protein activity, binding to essential
XX      PT      protein, by contacting protein with test compound and determining
XX      PT      effects -
XX      PS      Claim 1; Figure 79; 522pp; English.
XX      CC      The invention describes a method of screening a candidate antifungal
XX      CC      compound for interaction with essential proteins (EP) or for modulation
XX      CC      of EP activity e.g fungal gene transcription. The proteins tested in the
XX      CC      invention include RPC34, POP3, TFA2, NAB2, MPT1, MTR2, BOS1, POL30, RSA2,
XX      CC      SQT1, MTW1, TFB1, SPC98, BFR2, RNAL, GCD1, SKI6, NIP1, LCP5, NCE103,
XX      CC      ECO1, ORC2, CNS1, YPD1, TIM10 and SRB4 from S. cerevisiae; C. albicans
XX      CC      and human homologues. The method involves contacting a culture with one
XX      CC      or more test compounds and determining the effects on the growth or
XX      CC      viability of the culture of cells which preferably comprises fungal cells
XX      CC      or yeast cells. Preferably the identified compounds interact with, or
XX      CC      modulate (preferably inhibit) activity of C. albicans EP. The inhibitor
XX      CC      compounds identified by the method are useful for preventing or
XX      CC      inhibiting fungal, particularly C. albicans growth in culture or in a
XX      CC      mammal. The antifungal agents interact with essential fungal elements
XX      CC      that can be used to treat fungal infection by preventing the growth and
XX      CC      preferentially killing the fungi, but does not inhibit the biological
XX      CC      activity of mammalian homologues. This amino acid sequence represents a
XX      CC      target protein used to test the antifungal compounds, described in the
XX      CC      method of the invention.
XX      SQ      Sequence 533 AA;
XX      Query Match 5.6%; Score 127; DB 23; Length 533;

```

```

Best Local Similarity 21.4%; Pred. No. 0.11;
Matches 92; Conservative 61; Mismatches 157; Indels 120; Gaps 16;

Oy      5 IROFPIEYISXSGKQENRNGSI-----GPSYCKSIQNNQANSLDEPEGEPLDTLV 57
Db      9 ISDIAIKPVNSDFDIEDENASLFOHNEKNGES-----DSDYGNSTTEETKKAHYLEV 62
Oy      58 NRMORONTQOGDGVLDSTKK-TSISSESSICDPSSENSVAGRLHRRREDYVERSAEF- 115
Db      63 ----EKSLRAEKGLINDPKYTGKSGRQALYEVESENEDEEBEEDDEEEDALSTFR 118
Oy      116 -----ADG-----LTSKAL-----KDIQSGALDINKAGI 139
Db      119 TDSDEDEVEIDEEESDADGETEBAQKRHALSKLIQETKQAIKLSQSVGRDASKG-- 176
Oy      140 LYGIPQKTL-----LHLEALPAGKPASFKNKTDFHDSYSYKSKETCAVLQVALLWA 193
Db      177 -YSILOQTKLPDNIIDLRIRIKLOKAVIAANKPLTTESWEAKMDDEETKRLK----- 229
Oy      194 RAQARTESKUNLLETSEIKF-----PTASTYLHQLTIQKWTQPKENKESLQYETS 246
Db      230 --ENEKLFNNLFRNLINFRIRFQLGDHITONEEVAKHLSKRSKSLKELYQETNSLDSELK 287
Oy      247 N-PTVOLKIPOLRVSVSSKQPDGSGL-----LDWYQVSTSSVLEGSALQKXKN 296
Db      288 EYRTAVLKKWSTKVSASGMAALSNKFKAIINLPADVQVENQLSDMSRLMKRTKLN-RN 346
Oy      297 ILPKONKIECS-----GPTVHSSVDSYFLHGLSPLCLNSKNKGTGVTSENTEDGID 348
Db      347 ITPLFYQKDCANGRLPELISPVKDSVDD-----NENSDGDID 384
Oy      349 RKDSKQPRKX 358
Db      385 IPKNYDPRKX 394

RESULT 14
ABB63057
ID      ABB63057 standard; Protein; 2285 AA.
XX      AC      ABB63057;
XX      DT      26-MAR-2002 (first entry)
XX      DE      Drosophila melanogaster polypeptide SEQ ID NO 15963.
XX      KW      Drosophila; developmental biology; cell signalling; insecticide;
XX      KW      pharmaceutical.
XX      OS      Drosophila melanogaster.
XX      PN      WO200171042-A2.
XX      PD      27-SEP-2001.
XX      PF      23-MAR-2001; 2001WO-US09231.
XX      PR      23-MAR-2000; 2000US-191637P.
XX      PS      11-JUL-2000; 2000US-0614150.
XX      PA      (PEKE ) PE CORP NY.
XX      PI      Venter JC, Adams M, Li PWD, Myers EW;
XX      DR      WPI; 2001-656860/75.
XX      N-PSDB; ABL07160.
XX      PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      PT      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      PT      interactions -
XX      PS      Disclosure; SEQ ID NO 15963; 21pp + Sequence listing; English.
XX

```



Qy 361 RYROYDHEIMEEAIAMVMSGKMSVSKAQGIYGVPHSTLEYKVKERSGTLKTPPKKLRIP 420  
Db 1244 RTREQDVEVLBLEPLKCELVSGEST-----QNCEDRLPVGKTEANGKKKPSQOKTLEERP 1295  
Qy 421 -----DTGLYNYMTDSTGSGCKNSSK 440  
Db 1296 VNKCSDOIKLKNTTDDKNNENRESEK 1321

Search completed: October 28, 2003, 12:02:09  
Job time : 70.8626 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 : Search time 36.6101 Seconds  
(without alignments)  
510.826 Million cell updates/sec

Title: US-10-016-768a-8  
Perfect score: 2250  
Sequence: 1 MKKMIROPAIEYISKSGKTQ.....GLYNTDGGTSCSKNSKXPV 442

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	5.8	3248	1 US-08-353-700-1	Sequence 1, Appl
2	130	5.8	3248	5 PCT-US95-16216-1	Sequence 1, Appl
3	123	5.5	2482	1 US-08-328-254-6	Sequence 6, Appl
4	117	5.2	1146	4 US-08-914-999-6	Sequence 6, Appl
5	114.5	5.1	569	4 US-09-173-053-18	Sequence 18, Appl
6	114.5	5.1	1279	4 US-09-124-517-2	Sequence 2, Appl
7	114.5	5.1	1279	4 US-09-641-807A-2	Sequence 2, Appl
8	114.5	5.1	1279	4 US-09-723-096-2	Sequence 2, Appl
9	113.5	5.0	431	4 US-09-286-981B-3	Sequence 3, Appl
10	113	5.0	876	4 US-09-773-416-14	Sequence 14, Appl
11	111	4.9	1589	3 US-08-755-587-189	Sequence 189, App
12	111	4.9	1786	3 US-08-973-462-8	Sequence 8, Appl
13	110	4.9	535	2 US-08-007-107-2	Sequence 2, Appl
14	110	4.9	1196	4 US-09-107-532A-3944	Sequence 3944, Ap
15	110	4.9	1939	4 US-09-310-187A-1	Sequence 1, Appl
16	109.5	4.9	466	3 US-08-235-836C-107	Sequence 107, App
17	109	4.8	3878	4 US-09-914-259-11	Sequence 11, Appl
18	108.5	4.8	500	4 US-09-071-035-396	Sequence 396, App
19	108.5	4.8	1074	4 US-09-071-035-358	Sequence 358, App
20	108.5	4.8	1074	4 US-09-071-035-394	Sequence 394, App
21	108.5	4.8	2662	4 US-09-595-684B-31	Sequence 31, Appl
22	108.5	4.8	3696	4 US-09-134-001C-5080	Sequence 5080, Ap
23	108.5	4.8	8931	4 US-08-714-741-32	Sequence 32, Appl
24	108	4.8	630	3 US-08-973-462-9	Sequence 9, Appl
25	107.5	4.8	1087	4 US-09-914-259-12	Sequence 12, Appl
26	107.5	4.8	1792	2 US-08-962-284-4	Sequence 4, Appl
27	107	4.8	1404	4 US-08-801-308-1	Sequence 1, Appl

28	106.5	4.7	1588	5 PCT-US93-07261-11	Sequence 11, Appl
29	106.5	4.7	1663	5 PCT-US93-07261-16	Sequence 16, Appl
30	106.5	4.7	1664	1 US-09-599-652-2	Sequence 2, Appl
31	106.5	4.7	1664	2 US-08-642-846-2	Sequence 2, Appl
32	106.5	4.7	1664	4 US-09-264-604-2	Sequence 2, Appl
33	105.5	4.7	1541	3 US-08-296-791-3	Sequence 3, Appl
34	105.5	4.7	1541	5 PCT-US95-10661A-3	Sequence 3, Appl
35	105.5	4.7	1545	3 US-08-296-791-4	Sequence 4, Appl
36	105.5	4.7	1545	5 PCT-US95-10661A-4	Sequence 4, Appl
37	105	4.7	746	4 US-09-134-001C-3214	Sequence 3214, Ap
38	105	4.7	1312	4 US-09-345-882-29	Sequence 29, Appl
39	104.5	4.6	712	2 US-08-468-576B-17	Sequence 17, Appl
40	104.5	4.6	712	2 US-08-468-576B-17	Sequence 17, Appl
41	104.5	4.6	712	3 US-08-468-576B-17	Sequence 17, Appl
42	104	4.6	1040	3 US-08-961-083-118	Sequence 118, App
43	104	4.6	1040	4 US-09-536-784-118	Sequence 118, App
44	103.5	4.6	829	1 US-07-670-611-2	Sequence 2, Appl
45	103.5	4.6	829	1 US-08-220-674-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-353-700-1  
Sequence 1, Application US/08353700  
Patent No. 559919  
GENERAL INFORMATION:  
APPLICANT: YEN, TIMOTHY J.  
APPLICANT: RATTNER, JEROME B.  
TITLE OF INVENTION: NUCLEIC ACID ENCODING A  
TITLE OF INVENTION: TRANSLATION-EXPRESSED KINETOCORE PROTEIN,  
AND METHODS OF USE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN  
STREET: 1601 MARKET STREET, SUITE 720  
CITY: PHILADELPHIA  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0; Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,700  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: REED, JANET E.  
REGISTRATION NUMBER: 36,252  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3248 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: HUMAN  
US-08-353-700-1  
Query Match 5.8%; Score 130; DB 1; Length 3248;  
Best Local Similarity 20.2%; Pred. No. 0.028;  
Matches 104; Conservative 74; Mismatches 200; Indels 138; Gaps 21;

```

OY 7 QFAIEYISKSGKTOENR---NGSICPSIVCKSIOMNAENSLOEBOGCPDLDTVNRMOEQ 63
D 2676 QDTEVLQSSYKLNENLELTTRMDKMSFVEKYNMKTAKETELQREHMAQKTALEOEL 2735
OY 64 NTOOGDGVLDLSTKTSIKSEESSICDPSSENS-VAGRLHNRNEDYVERSAEFADGL--- 119
D 2736 SGEKNRLAGELQLLLEIKSSKDQKELTLENSLKKSLDCMHKDOVEKKGKVEBEIAEY 2795
OY 120 ---LSKALKDIOGALDINKAGILYGIPOKTLHLLEALPAGKPAEFKNTKTRDPHDSYS 176
D 2796 QLRHAEAEKKGQALLDPTNKQ---YEVEIQT-----YREKL----- 2828
OY 177 KDSKETCAVLQKVALMARQAERTKSKLN--LLETSEIKFPPTASTYHQLTLOKMTQF 234
D 2829 -TSKEECLSSOKLEI---DLKSKKEELNLSKATTQILELKTMDNL---KYVQL 2880
OY 235 KENKESL-----QYETSNPTVOLKIPOLRVSSVSKSQPDGSGLLDVWYQVSKT 282
D 2881 KKENERAQGMKMLIKSCQLEEEKEILOKELSQLQAQ-----EKQKT 2924
OY 283 SSYLE-----GSAIQKLNILPKONKIE-----CGPVTHSSVD--SYFLHGDLSPL 327
D 2925 GTVMDTKVDELTTTEIKELTLEKTKADEYLDKYCSLLISHEKLEKAKEMLETOVAHL 2984
OY 328 C-----LNSKNGTVDG-----TSENTEDEGLDRKDSKOPRK-----RGRYRQ 364
D 2985 CSQOSKQDSRSGPLGVPVPGSPPIPSYTEKRLSSGONKASGKORSSGIWENGRGTPA 3044
OY 365 YDHEIMEAIAWMSGKMSVSKAQGI---YGVPH-----STLEYKVKERSGT 408
D 3045 TPESFSKSKKAVWSGHPADTEGTEFEPEGLPEVVKKGFPADIPGKTSPYILRRTTMA 3104
OY 409 LKTPPK---KTLRLPDTGLYMTDSCGTSCKNSSKP 441
D 3105 TRTSPRLAQKLAISPSTL-----GKENLAESSKP 3134

RESULT 2
PCT-US95-16216-1
; Sequence 1, Application PC/TUS9516216
; GENERAL INFORMATION:
; APPLICANT: Yen, Timothy J.
; APPLICANT: Ratner, Jerome B.
; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently
; TITLE OF INVENTION: Expressed Kinetochore Protein, and Methods of Use
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16216
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,700
; FILING DATE: 09-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 3248 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-16216-1

Query Match 5.8%; Score 130; DB 5; Length 3248;
Best local Similarity 20.2%; Pred. No. 0.028;
Matches 104; Conservative 74; Mismatches 200; Indels 138; Gaps 21;

OY 7 QFAIEYISKSGKTOENR---NGSICPSIVCKSIOMNAENSLOEBOGCPDLDTVNRMOEQ 63
D 2676 QDTEVLQSSYKLNENLELTTRMDKMSFVEKYNMKTAKETELQREHMAQKTALEOEL 2735
OY 64 NTOOGDGVLDLSTKTSIKSEESSICDPSSENS-VAGRLHNRNEDYVERSAEFADGL--- 119
D 2736 SGEKNRLAGELQLLLEIKSSKDQKELTLENSLKKSLDCMHKDOVEKKGKVEBEIAEY 2795
OY 120 ---LSKALKDIOGALDINKAGILYGIPOKTLHLLEALPAGKPAEFKNTKTRDPHDSYS 176
D 2796 QLRHAEAEKKGQALLDPTNKQ---YEVEIQT-----YREKL----- 2828
OY 177 KDSKETCAVLQKVALMARQAERTKSKLN--LLETSEIKFPPTASTYHQLTLOKMTQF 234
D 2829 -TSKEECLSSOKLEI---DLKSKKEELNLSKATTQILELKTMDNL---KYVQL 2880
OY 283 SSYLE-----GSAIQKLNILPKONKIE-----CGPVTHSSVD--SYFLHGDLSPL 327
D 2925 GTVMDTKVDELTTTEIKELTLEKTKADEYLDKYCSLLISHEKLEKAKEMLETOVAHL 2984
OY 328 C-----LNSKNGTVDG-----TSENTEDEGLDRKDSKOPRK-----RGRYRQ 364
D 2985 CSQOSKQDSRSGPLGVPVPGSPPIPSYTEKRLSSGONKASGKORSSGIWENGRGTPA 3044
OY 365 YDHEIMEAIAWMSGKMSVSKAQGI---YGVPH-----STLEYKVKERSGT 408
D 3045 TPESFSKSKKAVWSGHPADTEGTEFEPEGLPEVVKKGFPADIPGKTSPYILRRTTMA 3104
OY 409 LKTPPK---KTLRLPDTGLYMTDSCGTSCKNSSKP 441
D 3105 TRTSPRLAQKLAISPSTL-----GKENLAESSKP 3134

RESULT 3
US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254

```

FILING DATE: 24-OCT-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/141,239  
 FILING DATE: 22-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Campbell, Cathryn A.  
 REGISTRATION NUMBER: 31,815  
 REFERENCE/DOCKET NUMBER: P-CJ 1191  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2482 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-328-254-6

Query Match 5.5%; Score 123; DB 1; Length 2482;  
 Best Local Similarity 20.0%; Pred. No. 0.082;  
 Matches 103; Conservative 74; Mismatches 201; Indels 138; Gaps 21;

7 QFAIEYISKGTQENR--NGSIGPSIVCKSIOMNOAENSIOEBOGPDLDITVNRMOQ 63  
 1948 QDTLEVLQSYKMLENELELTCKDKRSFYEKYNKTKAKTELOREHMAOKTAELOEEL 2007  
 64 NTOQDGVLDLSTKTKTSIKSESSICDPSSENS-VAGRLHRRNEDYVERSAEFADGL--- 119  
 2008 SEKRLAGELQLLEIEIKSKDQKELTEENSELKSSLDCKHKKDQVEKEGVREIARY 2067  
 120 ---LSKAKDIOGSLDINKAGILVGIPOKTLHLHLALPAGKPAKFKKTTDFHDSY 176  
 2068 QURLHEAEKQHALLDTNQ---YEVEIQT-----YREKL----- 2100  
 177 KDSKETCAVLQKVALMARQAERTKSKLN--LLETSEIKPTASTYHLQTLQKMTOP 234  
 2101 -TSKECGLSSQLEI---DLKSSKEELANSLKATQLEELKTKTMNL---KYVNL 2152  
 225 KENESL-----QYETSNPYQVKIPOLRVSSVSKSPDGGSLDVMYQVSKT 282  
 2153 KKENERAQGMKLLIKSCQLEEKELIQKEISQLQAAO-----EKQKT 2196  
 283 SSVLE-----GSAQKKNILPKONKIE-----CSGPVTHSSVD--SYFLHGDLSPL 327  
 2197 GTVMDTKVDELTEETKELEKTEKEADEYLDKXCSLLISHEKLEKAKEMILETOVAL 2256  
 328 C-----LNSKNGTVDC-----TSENTEDGLDRKDSKOPKX-----RGRYQ 364  
 2257 CSQOSKQDSRSGPLGPPVPGSPIPSVYTEKRLSSQNSASGKROKSSGIMWGGPTPA 2316  
 365 VDHEIMEEAIAMVMSGKMSVSKAOGI---YGVPH-----STLEKXKVERSGT 408  
 2317 TPESFSKSKKAVMGIGHPAEDTEGEFEPEGLPEVVKKGFADIPITGKTSPIILARTTVA 2376  
 409 LKTPK---KURLPDTGLYNNMTDSGTSCSKSKSP 441  
 2377 TRTSPLAOKLALSPSL-----GKENLABSKP 2406

RESULT 4  
 US-08-914-999-6  
 Sequence 6, Application US/08914999  
 Patent No. 6346406  
 GENERAL INFORMATION:  
 APPLICANT: Ryazanov, Alexey G.  
 APPLICANT: Hailt, William N.  
 APPLICANT: Pavut, Karen S.  
 TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.  
 STREET: 411 Hackensack Ave, Continental Plaza, 4th  
 STREET: Floor  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/914,999  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jackson Esq., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 601-1-078  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1146 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Dictyostelium discoideum  
 US-08-914-999-6

Query Match 5.2%; Score 117; DB 4; Length 1146;  
 Best Local Similarity 19.0%; Pred. No. 0.089;  
 Matches 94; Conservative 82; Mismatches 182; Indels 136; Gaps 20;

33 CKSIOMNAE--NSIOEBOGPDLDITVNRMOQ--NTOQDGVLDLSTKTKTSIKSESS 87  
 55 CSSFLVSKAEFPNHLKDDAQLQLAVKFKHQFDLHTQ---LMAHTEQWEDQLEKTM 110  
 88 ICDPSENSVAGRLHRRNEDYVERSAEF-----ADGLSKALDIOGSLDINK 136  
 111 KVRNHTDLSGAVQTKLDEIGKMAFAKVEQOQQLARLITQOIQEKSTSSPLVK 170  
 137 AGILVG-----IPKTLHLHLALPAGKPAKFKKTTDFHDSYK 178  
 171 GGISGGGGGGDDSDGANISMTSKQLOQLOSL-----SIKKKELTELSDELQKL 226  
 179 SKETCAVLQKVALMARQAERTK-SKLNLL-----ETSEIKPTASTYL----- 222  
 227 ERSTONIDIKI---KRIGEVNEKIDKQLVSTIDISIGKTKDSIGYLESIIKKVEK 283  
 223 ---HQLTLQKAVTQFKEX--NESLQYETSNPYQVKIPOL-----R 258  
 284 EKKSEQQLPDSKIESLKDKIKIETQOLDTSEVVKLLESTSSGNLWAGLNGTSGR 343  
 259 VSSVSKQPDGSGLL-----DVMYQVSKTSVLEGSALQKNIIPKONKIEGSPVT 311  
 344 PSSSHFIPSSVSAAMNINKNIEIMEEYKVEKLOKRIREDINTKAEKLSVENSVDN 403  
 312 HSSVDSYFLHGDLSPLCNSKNGTVDGTSENTEDGLDRKS-----KOPRKKRGY----- 362  
 404 RSEIEG-----LEKDCNGQPD-KQDNKIQVVEDLKKSDSILLMKNLKYVNEFVORE 456  
 363 -----ROYDHEIME--EAIAMVMSGKMSVSKAOGIYGVPHSTLE 399  
 457 RDSESERIKLQDSIKRLEONOKKILEAIOEGNEOVERVLRBASISP---ISSVKSPI- 512  
 400 YKVKERSGTLKTPP 413

Db 513 -TTKRSIIINSPP 525

## RESULT 5

US-09-173-053-18  
Sequence 18, Application US/09173053

Patent No. 6451769

GENERAL INFORMATION:

APPLICANT: HUEBNER, Robert C.

APPLICANT: NORMAN, Jon A.

APPLICANT: LIANG, Xiaowu

APPLICANT: CARMER, Kristin R.

APPLICANT: BARBOUR, Alan G.

APPLICANT: LUKE, Catherine J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ADMINISTERING BORRELIA DNA

FILE REFERENCE: 454312-2440.1

CURRENT APPLICATION NUMBER: US/09/173,053

CURRENT FILING DATE: 1998-10-15

PRIOR APPLICATION NUMBER: 08/663,998

PRIOR FILING DATE: 1996-06-14

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 18

LENGTH: 569

TYPE: PRT

ORGANISM: Borrelia burgdorferi

US-09-173-053-18

## Query Match

Best Local Similarity 5.1%; Score 114.5; DB 4; Length 569;

Matches 107; Conservative 92; Mismatches 176; Indels 169; Gaps 26;

Qy 25 GSIGSIYCK-SIQMNAENSL-----QEQGSPDL--TNRMQEQN-- 65  
Db 9 GLIALALACKQKNSVSLDKNSVSVLDPEEMKVLKSKKDKKYDLATVVKLEKGTSD 68  
Qy 66 -QOQGDVLDLSTKTKTSIKSESSICDPSESN-----SVAGRLHNR 105  
Db 69 KNSGSGVLE-----GVADSKSVKLTISDDLQGTTFLEFKEDKTLVSKVTSMDKST 122  
Qy 106 EDYVERSAEPFADGLSKA-----LKDTQSG-ALDINKAGIYG--IPQKTLHLLE 153  
Db 123 EEKNEKEVESEKITTIRADGTRLEVTGKSDSGKAKKEVLKGYVLEGTTLAEKTLVVE 182  
Qy 154 ALPAGKPAKFNKTRDFHDSYKDSKETCAVLQKVALMARQAERTKSKLNLETSI 213  
Db 183 ---GTVTLKNSIKSGEVSVELNDT-DSSATKTKTAAMNSGTSTLT--ITVNSKTKYDL 235  
Qy 214 KEPTASTYLHQLTLOQM-----VTOFKENKESLOY----- 243  
Db 236 VFTKENT-----ITVQYDSNGTKLEGSVEITKDEIKNALKMLLIGFALALALIGCAQ 291  
Qy 244 -----ETSNPTVQLKIPOLRVSSVS-----KSQPDGSGLLD 274  
Db 292 KGAESIGSQKENDLNLIEDSSKSHQNAKODLPATVEDSVSLFNGNKIFVSEKNSGKYD 351  
Qy 275 VMYQVSKT-----SSVLBESALQKLNILPKONKIECGSPVTHSSVDSTFLHGD 323  
Db 352 LRATIDQVELKGTSDKNNGSGTLBEGSKPKSK-----VKLTVSAIDLNTVTLLEAF---D 401  
Qy 324 LSPCLNKNKNGTVDGTSENTEDGL--DRKSKOPKPKGRROYDH--EIMEEATAMVWSG 380  
Db 402 ASNQKISSKVTKKQOSI--TEETLKANKLDSKKLTRSGTTLSEISQITDADNATAVETL 459  
Qy 381 KMVSYSKAGIYGVPHSTLEYKVKERSGTLKTPPKK---KLRLPDGTGLYNNMDSGTSGCK 436  
Db 460 KNSI-KLEGSIVGKTVIE--IKEGTVTLKREIEKDGKVKYFLNDT-----AGSNK 507  
Qy 437 NSSK 440  
Db 508 KTGK 511

## RESULT 6

US-09-724-517-2

Sequence 2, Application US/09724517

Patent No. 6379941

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe

APPLICANT: Freedman, Richard

TITLE OF INVENTION: No. 6379941el motor proteins and methods for

their use

FILE REFERENCE: 1031

CURRENT APPLICATION NUMBER: US/09/724,517

CURRENT FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: US/09/641,807

PRIOR FILING DATE: 2000-08-17

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 1279

TYPE: PRT

ORGANISM: Human

FEATURE:

NAME/KEY: VARIANT

LOCATION: (409)...(436)

OTHER INFORMATION: Xaa = any amino acid

US-09-724-517-2

## Query Match

Best Local Similarity 5.1%; Score 114.5; DB 4; Length 1279;

Matches 86; Conservative 75; Mismatches 153; Indels 133; Gaps 18;

Qy 15 KSGKTOENRNGSIPIYSICKSIQMNQAEENSLQEEQEGPLDLYVRMOEQNTQOQDGVLDL 74  
Db 534 KSGTRCKSRSMIQKQPDVCSVLSPDTQDQKSDLENELKIDCLOESQE-----LNL 587  
Qy 75 STKTKTSIKSESSICDPSSNSVAGRLHNRREDYVERSAEPADGLSKALKDIOGALDI 134  
Db 588 QKLKNS-----ERILTEAKQKRE--LTI 609  
Qy 135 NKAGILVIGPQKTLHLLEALPAGKPAKFNKTRDFHDSYKDSKETCAVLQKVALMAR 194  
Db 610 N-----IKMKEDLIK-ELIKTGNDKSVSK-----QYSLKVTK-----LEHDA--- 646  
Qy 195 AQARTKSKLNLEL-----SEIKFPTASTYLHQLTLOKQVTOFKENKESLOYETSNPTV 250  
Db 647 -----EQAKVELLETQKQOLENKLSDVAMKVKQK--ERRKMDA----- 687  
Qy 251 QLKIPOLRVSSVSQKSPDGSGLLDWMYQVSKTSSVLEGSA---LQK--LKNILPKONKI 304  
Db 688 ---AKLRVQVLQKQODSKKLASLQIENKRALEQSVDHMKYQKIQOLQKLEENE- 742  
Qy 305 ECGSPVTHSSVDSTFLHGDLSPLCLNKNKNGTVDGTSENTEDGLDRKDSKOPKPKGRYRQ 364  
Db 743 ---KRDLDVAIKEDQKIKIEIQKGTQEGQLKPAKD---LDACNLKRRKSGFGS 792  
Qy 365 YDH-----EIMEEATAMVWSGKMSVSKAGIYGVPHSTLEYKVKERSGTLKTPP--- 413  
Db 793 IDHQLKDEQKMLDEVEKVLNROGLEE-----LEADLKKRAIYSKEALL 841  
Qy 414 KKKLRLPDGTGLYNNMDSGTSGCKNSK 440  
Db 842 QEKSHLENKKLRSQALNTDLSLKISTR 868

RESULT 7  
US-09-641-807A-2  
Sequence 2, Application US/09641807A  
Patent No. 6440731  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Freedman, Richard  
TITLE OF INVENTION: No. 6440731el motor proteins and methods for  
their use  
FILE REFERENCE: 1031

CURRENT APPLICATION NUMBER: US/09/641,807A  
CURRENT FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1279  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (409)...(446)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-641-807A-2

Query Match 5.1%; Score 114.5; DB 4; Length 1279;  
Best Local Similarity 19.2%; Pred. No. 0.18;  
Matches 86; Conservative 75; Mismatches 153; Indels 133; Gaps 18;

QY 15 KSGKTQENRNGSIGPSIVCKSIQNMQAENSLQEEQEGPLDLTVNMQEONTQGGDGLDL 74  
DB 534 KSGTRCRSRSMIQRKDSVSLVELSDTQDETQKSLDLENDDKIDCLQSSQE-----LNL 587  
QY 75 STKTSISESSSICDPSESSVAGRLHNRNEDYVERSAEPADGLLSKALKDIOGALDI 134  
DB 588 QKLKNS-----ERILTEAKQKMR--LTI 609  
QY 135 NKAGILYGIPOKTLHLLEALPAGKPAFKNKTDFHDSYKDSKETCAVLQKVALMAR 194  
DB 610 N-----IKMKEDLIK-ELIKTGNDAKSVK-----QYSLKVTK-----LEHDA----- 646  
QY 195 AQERTKSKNLLET-----SEIKPTASTYLHQLTLQKMTQPFKEKNESLQYETSNTPTV 250  
DB 647 -----EQAKVELIETQKQLEENKDLSDVAMKVKLQK--EPKKMDA----- 687  
QY 251 QLKIPOLRVSSVSKSQPDGSGLLDVMYQVSKTSVLEGSA-----LQK--LKNILPKONKI 304  
DB 688 -----AKLRVQVLQKKQDSKTLASLQNEKRANELEQSVDMKVKYQKIQLOKRLRENE- 742  
QY 305 ECGSVTHSSVDSYFLHGDLSPLCLNSKNGTVDGTSENTEDELDRKDSKQPKKRGYRQ 364  
DB 743 -----KKQQLDAVIRKDOQKIKEIQLKTGOEGGLKPAED-----LDACNLKRRKSGFS 792  
QY 365 YDH-----EIMEEALAMVMSGKMSVSKAGIYGVPHSTLEKVKERSGTLKTPP--- 413  
DB 793 IDHLQKLEQKKWLEDEVEKVLNQOELE-----LEADLKREALIVSKKALL 841  
QY 414 KKKLRLPTGLYNTDSCGTGCKNSK 440  
DB 842 QEKSHLENKKLRSQALNTDSLKISTR 868

RESULT 8  
US-09-723-096-2  
Sequence 2, Application US/09723096  
Patent No. 6448026  
GENERAL INFORMATION:  
APPLICANT: Berard, Christophe  
TITLE OF INVENTION: No. 6448026el motor proteins and methods for  
FILE REFERENCE: 1031  
CURRENT APPLICATION NUMBER: US/09/723,096  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: US/09/641,807  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1279  
TYPE: PRT  
ORGANISM: Human  
FEATURE:  
NAME/KEY: VARIANT

LOCATION: (409)...(436)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-723-096-2

Query Match 5.1%; Score 114.5; DB 4; Length 1279;  
Best Local Similarity 19.2%; Pred. No. 0.18;  
Matches 86; Conservative 75; Mismatches 153; Indels 133; Gaps 18;

QY 15 KSGKTQENRNGSIGPSIVCKSIQNMQAENSLQEEQEGPLDLTVNMQEONTQGGDGLDL 74  
DB 534 KSGTRCRSRSMIQRKDSVSLVELSDTQDETQKSLDLENDDKIDCLQSSQE-----LNL 587  
QY 75 STKTSISESSSICDPSESSVAGRLHNRNEDYVERSAEPADGLLSKALKDIOGALDI 134  
DB 588 QKLKNS-----ERILTEAKQKMR--LTI 609  
QY 135 NKAGILYGIPOKTLHLLEALPAGKPAFKNKTDFHDSYKDSKETCAVLQKVALMAR 194  
DB 610 N-----IKMKEDLIK-ELIKTGNDAKSVK-----QYSLKVTK-----LEHDA----- 646  
QY 195 AQERTKSKNLLET-----SEIKPTASTYLHQLTLQKMTQPFKEKNESLQYETSNTPTV 250  
DB 647 -----EQAKVELIETQKQLEENKDLSDVAMKVKLQK--EPKKMDA----- 687  
QY 251 QLKIPOLRVSSVSKSQPDGSGLLDVMYQVSKTSVLEGSA-----LQK--LKNILPKONKI 304  
DB 688 -----AKLRVQVLQKKQDSKTLASLQNEKRANELEQSVDMKVKYQKIQLOKRLRENE- 742  
QY 305 ECGSVTHSSVDSYFLHGDLSPLCLNSKNGTVDGTSENTEDELDRKDSKQPKKRGYRQ 364  
DB 743 -----KKQQLDAVIRKDOQKIKEIQLKTGOEGGLKPAED-----LDACNLKRRKSGFS 792  
QY 365 YDH-----EIMEEALAMVMSGKMSVSKAGIYGVPHSTLEKVKERSGTLKTPP--- 413  
DB 793 IDHLQKLEQKKWLEDEVEKVLNQOELE-----LEADLKREALIVSKKALL 841  
QY 414 KKKLRLPTGLYNTDSCGTGCKNSK 440  
DB 842 QEKSHLENKKLRSQALNTDSLKISTR 868

RESULT 9  
US-09-286-981B-3  
Sequence 3, Application US/09286981B  
Patent No. 6503511  
GENERAL INFORMATION:  
APPLICANT: Wizemann, Theresa M.  
APPLICANT: Koenig, Scott  
TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines  
FILE REFERENCE: 469201-396  
CURRENT APPLICATION NUMBER: US/09/286,981B  
PRIOR FILING DATE: 1999-04-06  
PRIOR APPLICATION NUMBER: US 60/085,743  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-286-981B-3

Query Match 5.0%; Score 113.5; DB 4; Length 431;  
Best Local Similarity 19.2%; Pred. No. 0.041;  
Matches 57; Conservative 55; Mismatches 118; Indels 67; Gaps 7;

QY 10 IEYISKGTQENRNGSIGPSIVCKSIQNMQAENSLQEEQEGPLDLTVNMQEONTQGGD 69  
DB 111 VEEAEKAKQKQKEBHRNYPITTYKTLELTAESDVEYK--AELELVKEAKGSRNEE 167  
QY 70 GVLDSLTKTSIKSESSICDPSESSVAGRLHNRNEDYVERSAEPADGLLSKALKDIOG 129

Db 168 KIKKAAVESKKAATKLEIKTERKKA-----EEAKRKAEEVEVKNLKKRRKR 220  
Qy 130 GADINKAGIIVGIPQKTLHLHLALPAGKAPSEFNKTRDPHDSY-----YKDS 179  
Db 221 GAF-----GEPATPKKENDAKSSPSSVYKSSKPKLKSE 255  
Qy 180 KETCAVLQKVALMARAERTKSKTLNLESEIPEFPASTYVHQLTLQKAVTOPEKEN- 238  
Db 256 KKVAAEKVAAEAKVAAEAKKAK-DQKEDRRRYPPNTYKTELELEIAESDVAKKEAL 314  
Qy 239 ESLOYETSNPTVOLKIPOLRVSSVSKSQPDGGLDVMYQVSKTSSVLEGSALQK 295  
Db 315 ELVKEAPEQONBEKIKQAKKVESK-----AEATRLKIK 351

RESULT 10  
US-09-773-416-14  
Sequence 14, Application US/09773416  
Patent No. 6483725  
GENERAL INFORMATION:  
APPLICANT: No. 6483725eborn, Mathieu  
APPLICANT: Damen-van Oorschot, Astrid  
APPLICANT: Rohm, Jennifer  
APPLICANT: Weiss, Berttram  
APPLICANT: Toschi, Luisaella  
TITLE OF INVENTION: Apoptin-associating protein  
FILE REFERENCE: 2906-4997US  
CURRENT APPLICATION NUMBER: US/09/773,416  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: PCT/NL00/00905  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: EP 99204242.4  
PRIOR FILING DATE: 1999-12-10  
PRIOR APPLICATION NUMBER: EP 00250119.5  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO. 14  
LENGTH: 876  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-773-416-14

Query Match 5.0%; Score 113; DB 4; Length 876;  
Best Local Similarity 20.5%; Pred. No. 0.14;  
Matches 102; Conservative 69; Mismatches 177; Indels 150; Gaps 24;

Qy 13 ISKSGKTQENRNGSIGPSIVCKSIQNOAENSLOEEOGPDLDITVNRMOEQNTQOGDVL 72  
Db 352 LKKIGDSSKNSDS-----QSVSSNTDADDTQEKNA-----TSNRKSSVGVK 393  
Qy 73 DLSTKTKISKEESSICDPSSENSVAGRL-HNRREDYVERSAEFPDLISK----- 122  
Db 394 KNSKSRITLROSMRI--PASNSSTSSKLTTHNNSRVPRKLLKPAKPLSKIKLHNCKR 451  
Qy 123 -----ALMDIOSGALDINKAG-ILYGIPOKTLHLHLALPAG-----KPAQFK 164  
Db 452 LEQKVASRLEWGNVLKEPKVLY---KNLPKDKPEBPQAQAAVAGCLTTHARE 507  
Qy 165 NKTDRPHDSYKDSKETCAVLQKVALMARAQERTKSKTLNLESEIKFP--TASYVL 222  
Db 508 HRQNVRAHSGES-SECTYITR-----RSVTRITNLKASASDIKLEPNTLNGVK 556  
Qy 223 HOLTLQKAVTOPEKENESLOYETSNPTVOLKIPOLRVSSVSKSQPDGGLDVMYQVSKT 282  
Db 557 SSVT-----EPCPDGGEQIQ-----PAPVLOEEELAHETAQGE-----AKC 593  
Qy 283 SSVLEGSALQK--NILPKNKIKESGCVTHSS-----VDSYFLHGDLSPLCLSKN 333  
Db 594 HKSDTGMSKKRSQKQKLVQFAKIEESTPV-HDSQKDDAVDLWGPFSH-----QGEHS 647  
Qy 334 GTVD-----GTSENTDGLDRKDS--KQPRKRGVRYQYHEIMEEAIAMVM 378

Db 648 GTVGVSVYTDCAAPSIVGCVTSVSDSFKTDSPRTAKSKKKRITRYDAQLIENNS--- 704  
Qy 379 SGKSVSKAGCIYGVPHSTL-----EYKVERSGTLKTPPKKALRP-----DTGILY-- 425  
Db 705 -----GIPXLTIRRRHDSSSKTDNDGNMSSKISIKLXKDHDNDNMLYVA 751  
Qy 426 ---NMTDGTGSCNKSX 440  
Db 752 KLNNGFNSGSGSSSTKLK 769

RESULT 11  
US-08-755-587-189  
Sequence 189, Application US/08755587  
Patent No. 6045997  
GENERAL INFORMATION:  
APPLICANT: Futreal, Phillip A  
APPLICANT: Wooster, Richard F  
APPLICANT: Ashworth, Alan  
APPLICANT: Stratton, Michael R  
TITLE OF INVENTION: Materials and methods relating to the  
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer  
TITLE OF INVENTION: susceptibility gene and uses thereof.  
NUMBER OF SEQUENCES: 222  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell Seltzer Park & Gibson  
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107  
CITY: Raleigh  
STATE: NC  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/755,587  
FILING DATE: 25-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9523959.6  
FILING DATE: 23-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525555.0  
FILING DATE: 14-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9617961.9  
FILING DATE: 28-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenneth D Sibley  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5405-135  
INFORMATION FOR SEQ ID NO: 189:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1589 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-755-587-189

Query Match 4.9%; Score 111; DB 3; Length 1589;  
Best Local Similarity 19.0%; Pred. No. 0.53;  
Matches 80; Conservative 56; Mismatches 103; Indels 182; Gaps 18;

Qy 56 TVNRMOEQNTQOGDVLDTKTKTSIKSESSICDPSSENSVAGR---LHNRREDYVE- 110  
Db 620 SVFPIKQKTEKSD-----EKSSKQVTLQNNIEWTCIFGVRNPEKYIKN 665  
Qy 111 -----RAEFADGILSKALDIOGALDINKAGILVGIPOKTLHLHLALPAG 158  
Db 666 TKHEDSYTSQRNKLNSDGSWST---SGPYIHKGD-----SDLPAD 706  
Qy 159 K---PASFNKTRDPHDSYKDSKETCAVLQKVALMARAQER---TEKSKTLNLETS 211  
Db 707 QGSKCPESCIOYAAEENTQIKENISDLTCLIMKAERTCMSSDKKQLPSPDMEQNIKEFN 766

QY 212 EIKFPTA-----STYHQLTQKVVTOFKEX 237  
 DB 767 -ISFOTASGKTRVSKESLUNKSVINFNRETDELTVISDSLSKILHGIKDKKHTSCHK 825  
 QY 238 NESL-----OYE--TSNPVT-----OLKIPOLRVSSVSK 264  
 DB 826 AISIKVFPEDHPITVVSQLPAAOHPPEYEIESTKEPTLSFHTPASKVKVIMQESLDKV-- 883  
 QY 265 SQPDGGLDWMYOVSKTSSVLEGS-----ALQKLK----- 295  
 DB 884 -----KNLFEFOYAKRTASFGSGSKPLKDSKKELTAYEKEIETVASKCEMKNFVSKET 938  
 QY 296 NILPKONKIEGSGPVTHSVDSYFLHGDLSPLCLNSKNGTGVDSNTEDGLDRKDSKOP 355  
 DB 939 EMLPQON-----YHMYROTEN-LKTSNGTSSKVOENIENNV-----KNP 977  
 QY 356 R 356  
 DB 978 R 978

RESULT 12  
 US-08-973-462-8  
 ; Sequence 8, Application US/08973462B  
 ; Patent No. 6191270  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DRUILHE, PIERRE  
 ; APPLICANT: DUBBERSIES, PIERRE  
 ; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
 ; FILE REFERENCE: 0660-0125-0 PCT  
 ; CURRENT APPLICATION NUMBER: US/08/973,462B  
 ; EARLIER FILING DATE: 1998-02-06  
 ; EARLIER APPLICATION NUMBER: PCT/FR96/00894  
 ; EARLIER FILING DATE: 1996-06-12  
 ; EARLIER APPLICATION NUMBER: FR 95/07007  
 ; EARLIER FILING DATE: 1995-06-13  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1786  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide  
 US-08-973-462-8

Query Match 4.9%; Score 111; DB 3; Length 1786;  
 Best Local Similarity 21.6%; Pred. No. 0.63;  
 Matches 72; Conservative 62; Mismatches 135; Indels 64; Gaps 15;

QY 26 SIGPSIVCSIQMNAENSLQEOEGPLDLTVNRMOEONTQCGDGLDLSTKTSIKSEE 85  
 DB 750 SVAPSYVE--ESVEEN--VEESVAENVEESVAENVEESVAENVEE-----SVAPTV 795  
 QY 86 SSICDSSSENSVAGRLHNRREDYVERSAFADGLSKALKDIOGSLDINKAGILYGIQ 145  
 DB 796 EETVAPSVESVAESVAENV--ATNLSDELNLGLGLETEI---KQSLNLEIE 850  
 QY 146 -----KTLHLLEALPAGKSPAFKNTKTRDPHDSYKDS--KETCAVLQKALMARAQAE 198  
 DB 851 VKNENVTTLLENVEETJASVTTFNSILBEIOENTTNTDIEKLEELHENVLSALENT 910  
 QY 199 RIEKSKLNLLET--SEIKFPTASTYVHLTQKVVTOFKESKNSLOYETSNPTVOLKIPOL 257  
 DB 911 OSEEBEKEVIVDIEVEKEEVAT-----TLIEFVEQAEBSKAN-----TITTEIFENL 956  
 QY 258 RVSSVSKQPDGSGLDWMYOVSKTSSVLEGSALQKLKNIILPKONKIEGSG--PVTHSSVD 316  
 DB 957 EENAVESNE-----NVAKENLEKLETVFNTVLDKV-----EETVEISGSLENNEMD 1003  
 QY 317 SYFLHGDLSPLCLNSKNGTGVDSNTEDGLDR 349

DB 1004 KAF-----FSEIFDN-----VKGIOENLLTGMR 1027

RESULT 13  
 US-08-007-107-2  
 ; Sequence 2, Application US/08007107  
 ; Patent No. 5837545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Guy, Charles L.  
 ; APPLICANT: Haskell, Dale W.  
 ; APPLICANT: Hoffig, Andrea  
 ; APPLICANT: Neven, Lisa  
 ; TITLE OF INVENTION: No. 5837545el Genes, Polypeptides, and  
 ; TITLE OF INVENTION: Compositions for Cold Tolerance and Drought Resistance in  
 ; TITLE OF INVENTION: Plants  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Saliwanchik & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: USA  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/007,107  
 ; FILING DATE: 19930121  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Saliwanchik, David R.  
 ; REGISTRATION NUMBER: 31,794  
 ; REFERENCE/DOCKET NUMBER: UF/S&S-109  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 904-375-8100  
 ; TELEFAX: 904-372-5800  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 535 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-007-107-2

Query Match 4.9%; Score 110; DB 2; Length 535;  
 Best Local Similarity 19.7%; Pred. No. 0.12;  
 Matches 97; Conservative 69; Mismatches 180; Indels 146; Gaps 21;

QY 20 QENRNGSIGPSIVCSIQMNAENSLQEOEGPL--DLTVNRMO--EONTQCGDGLDLSTR 77  
 DB 96 EONKGGVFG-KIKEKLPQOHOSDTHHTQQLVPASDHNYNTHNVHNODEKKNIILDKKD 154  
 QY 78 KTSIKSESSISIDPSSENSVAGRLHNRREDYVERSAFADGLSKALKDIOGSLDINKA 137  
 DB 155 KLPQGHED-----KKQDYHQHOEBEKKGGLDKIXDKLPQGNAGHT 196  
 QY 138 GILYGIPOKTLILL-----LEALPAGKSPAFKNTKTRDPHDSYKDSKETCAV 185  
 DB 197 QOQYAPRPHNYNTHNVHNODEENKDSVLDKIKDLFGOHEDKKNYH--HHOEBEKDSV 253  
 QY 186 LQKVALMARAQAE-----RTEKSKLNLLETSEIKFP--TASTYVHLTQKVM-- 230  
 DB 254 LDKIKDKKSGCHEDKKNYHNNHQBEEKGGVLDKIKDLPGOHDAVDYRHTQQLVPADH 313  
 QY 231 -----VTOFKESKNSLOYETSNPTVOLKIPOLRVSSVSKQPD-----SGGLLD 274  
 DB 314 NYNTHNVHNODEENKDSVL-----DKIKDLPGOHDDKAAYSOHNDHNNKHNOEBENKGVLD 368  
 QY 275 V-----MYQVSKTSSVL-----EGSALQKLKNIILPKO 301

Db 369 KIKDLPGVYVWVKHDGDIHEHTQQLYPAPDHNYNTHYVHEDEKKDSVLDKIKXKLPQ 428  
 QY 302 NIEESGPTHTSSVSVY---FLHGDSPLCLNSKNGTVGTEBNTEDGIDRDSKQPRK 358  
 Db 429 HE-EKAAVSEPSYSHPTPAKHHDYFPQEEKKGGMKID-----KLSQOHKXK 479  
 QY 359 RGRYQYDHEIMEEAIAMWSGMSVSKAQIGYVPHSTLEYKVERSGTLTKPPKKLR 418  
 Db 480 AD-----EHELVAPLVTV-----EPHSGD---KEKGFLE---KIKDX 512  
 QY 419 LPDTGLYNNFTDS 430  
 Db 513 IP--GLHSKXDA 522

RESULT 14  
 US-09-107-532A-3944  
 ; Sequence 3944, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:

APPLICANT: LYNN A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354

COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: PC  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,532A  
 FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-8277  
 TELEFAX: (781)893-5007

INFORMATION FOR SEQ ID NO: 3944:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1196 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium

FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (B) LOCATION 1...1196  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3944:  
 US-09-107-532A-3944

Query Match 4.9%; Score 110; DB 4; Length 1196;  
 Best Local Similarity 21.6%; Pred. No. 0.42;  
 Matches 93; Conservative 69; Mismatches 166; Indels 102; Gaps 22;  
 5 IRQAFIEYISKSGTKQENRNGSIGSIYCKSIQNN-QAENSLOEBOEGPLDLTVNRMOEQ 63

Db 251 LAKFNLELGKLSSESIQOE-----SLAKQKKNQAQDRLLIKNQVLLDLSEKLO-- 302  
 QY 64 NTQOGDGLDLSTKTKTSIKSESSICDPSSNSVAGRLHRNRREDYERSAFADGLSKA 123  
 Db 303 -TEGQKDLQBRTHGTOKSSQEQYTSIAEAKKVK-HFEKLQESLMKAAAE-KETEIOKA 359  
 QY 124 LKDIOGALDINKKGIYVGIPOKTLHLHLALPAGKPSFNGKTRDPFDHSYVD-SKET 182  
 Db 360 EANLKTQOELEK-----YKSTKELLAE-----RD-----QYVDMQEQ 395  
 QY 183 CAVLOKVALMARQAERTKSKNLLETSEIKFPPTASTYLH-----QLTLQKRVTFK 235  
 Db 396 AAVGNELKYLBRQYIOETAKSKQTLAKQSEVEASVDRLMKQEBLTOQAQLKSLTBTK 455  
 QY 236 EKNSLOQYETNPVQVKIPOLRVSVSKSPDSSGLDVIYQVSKTSVLEG--SALQK 293  
 Db 456 EKLEWIDQNGK-----KFOE--ALAKEQK-----MYOLMNOVOQLRARKKSIQ 498  
 QY 294 LKN-----ILPKNKIEGSPVTHSSVDSYFLHGDSPLCLNSKNGTVDTSEN-- 342  
 Db 499 IQENVFGYQGVRLVLQHKQQLSG-IVGAVAEILDVPADEF-LAIEALG--GAQGHVI 553  
 QY 343 TEDGLDRKDS---KQPKKRGY-----ROYDHEIMEEAIAMWSGMSVSKAQGI 390  
 Db 554 VENEDARQAITYLKQQRGRATFLPTTIKPRQLPAHILTOAAV-----EGF 602  
 QY 391 YGVPHSTLEY 400  
 Db 603 IGIASEOVSY 612

RESULT 15  
 US-09-310-187A-1  
 ; Sequence 1, Application US/09310187A  
 ; Patent No. 6358751  
 ; GENERAL INFORMATION:

APPLICANT: Benichou, Gilles  
 APPLICANT: Fedoseyeva, Eugenia  
 TITLE OF INVENTION: Involvement of Autoantigens in Cardiac  
 TITLE OF INVENTION: Graft Rejection  
 FILE REFERENCE: UCSF-090  
 CURRENT APPLICATION NUMBER: US/09/310,187A  
 CURRENT FILING DATE: 1999-05-12  
 NUMBER OF SEQ ID NOS: 3  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 1  
 LENGTH: 1939  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-310-187A-1

Query Match 4.9%; Score 110; DB 4; Length 1939;  
 Best Local Similarity 18.7%; Pred. No. 0.89;  
 Matches 92; Conservative 82; Mismatches 172; Indels 146; Gaps 19;

QY 18 KTOENRNGSIGPSIVCKSIQNNQAENSLOEBOEGPLDLTVNRMOEQTOGD-----GV 71  
 Db 1013 QVEDKYNLSKSKVKLEQVVDLEGSLEQKKYRMLF---RAKKLEGDCLKTQESI 1068  
 QY 72 LDLSTKK---TSIKSESSICDPSS---ENSVAGRLHRNRREDYERSAFADGL---- 119  
 Db 1069 MDLENDRQLQLEKKKKKEPDIQNSKIETDEQALALDLOKKKLENQARIIELEBELAER 1128  
 QY 120 -----LSKALKDIO-----SCALDINKAGILYGIPOKTLHLHLALPA 157  
 Db 1129 TARAKVEKLRSDLSRELEBISERLEBAGATSVQIENMKK----- 1168  
 QY 158 GKPSFKNKTRDPFDHSYVDSKETCAVLQK-----VA-----LMARQAERTE 201  
 Db 1169 -REAFQKMRDLEBATLQHEA--TAAALKRKIADSVAEIGEODINIQRYKQKLEKEKSE 1225  
 QY 202 -----SKNLLETSEIKFPPTASTYLHQL-TLOKWTQFKEKNESIQYE 244

```

Db      1226 FPLELDVTSNMEQIIKAKANLEKVSRTLEDQANEYRVKLEBAQBSLANDFTTORAKLQTE 1285
Qy      245 TSNPTVQLKI POLRVSSVSXSQPDSSGLDVMVQVSKTSSYLE-----GSAIOXKX 295
Db      1286 NGELARQLEBEKEALISQLTR-----GKLSYTOQMEDLKRQLEEBGKAKNALAHALQSA 1339
Qy      296 ---NILPKONKIECSPVTHSSVDSYFLHGDLSPLCLNSKNGTVDSGTSENTEDGLDRKDS 352
Db      1340 HDODLREQYEETEAE-----KAELORVLSKANSEVAQWRTKYETDAIQRTTEE 1387
Qy      353 KOPRKRGRYROYDHEIMEBAIAMWSGKMSVSKAQGIYGVPHSTLEYVK-----ERS 406
Db      1388 LEEAKKKLAQRLQD---AEEAVEAVNAKCSLEKTK-----HRLQNEIEDIMVVERS 1437
Qy      407 GTLKTPPKKKLR 418
Db      1438 NAAAAALDKKOR 1449

```

Search completed: October 28, 2003, 12:05:11  
 Job time : 40.6101 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:03:24 ; Search time 82.4495 Seconds  
(without alignments)  
901.011 Million cell updates/sec

Title: US-10-016-768A-8  
Perfect score: 2250  
Sequence: 1 MKKMIROFAIEYISKSGKTQ.....GLYNTDSCGCKSSKPV 442

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 629382 seqs, 167460630 residues  
Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2250	100.0	442	14	US-10-016-768-8
2	1411	62.7	277	12	US-10-029-386-33895
3	273	12.1	53	14	US-10-016-768-2
4	233.5	10.4	54	14	US-10-016-768-3
5	229	10.2	53	14	US-10-016-768-4
6	200.5	8.9	1165	14	US-10-016-768-10
7	165	7.3	53	14	US-10-016-768-1
8	163	7.2	53	14	US-10-016-768-5
9	131.5	5.9	848	12	US-10-011-588-45
10	131.5	5.8	870	14	US-10-029-386-32827
11	127	5.6	534	11	US-09-893-519A-37
12	124.5	5.5	972	10	US-09-924-154-16
13	121.5	5.4	1743	12	US-09-882-227-624
14	120	5.3	2696	12	US-10-309-933-4
15	119.5	5.3	1240	12	US-10-032-585-7319

16	117.5	5.2	1225	15	US-10-177-293-332	Sequence 332, App
17	117	5.2	1146	10	US-09-994-485-6	Sequence 6, Appl
18	117	5.2	1940	12	US-09-738-630-99	Sequence 99, Appl
19	116.5	5.2	1239	12	US-10-124-805-577	Sequence 577, App
20	116.5	5.2	1239	14	US-10-007-805-577	Sequence 577, App
21	116.5	5.2	1239	15	US-10-076-622-577	Sequence 32, Appl
22	115	5.1	877	12	US-10-309-422-32	Sequence 32, Appl
23	115	5.1	926	12	US-10-309-422-36	Sequence 36, Appl
24	115	5.1	961	12	US-10-309-422-40	Sequence 40, Appl
25	115	5.1	1042	12	US-10-309-422-40	Sequence 40, Appl
26	115	5.1	1043	12	US-10-309-422-20	Sequence 20, Appl
27	115	5.1	1091	12	US-10-309-422-12	Sequence 12, Appl
28	115	5.1	1092	12	US-10-309-422-24	Sequence 24, Appl
29	115	5.1	1126	12	US-10-309-422-18	Sequence 18, Appl
30	115	5.1	1127	12	US-10-309-422-28	Sequence 28, Appl
31	115	5.1	1146	10	US-09-832-292-10	Sequence 10, Appl
32	114.5	5.1	1952	11	US-09-840-743-17	Sequence 17, Appl
33	114.5	5.1	2476	12	US-09-824-574-7	Sequence 7, Appl
34	114	5.1	2344	9	US-09-815-242-12713	Sequence 12713, A
35	113.5	5.0	431	12	US-10-254-995-3	Sequence 3, Appl
36	113.5	5.0	1639	15	US-10-087-664-10	Sequence 10, Appl
37	113.5	5.0	2283	12	US-10-172-502-4	Sequence 4, Appl
38	113	5.0	876	9	US-09-773-416-14	Sequence 14, Appl
39	113	5.0	876	10	US-09-764-176-10	Sequence 10, Appl
40	113	5.0	876	11	US-09-733-416A-14	Sequence 14, Appl
41	113	5.0	1531	12	US-10-247-671-156	Sequence 156, App
42	113	5.0	5795	9	US-08-815-242-12610	Sequence 12610, A
43	111	4.9	1786	10	US-09-742-096-3	Sequence 6, Appl
44	110	4.9	1007	9	US-09-783-320-6	Sequence 6, Appl
45	110	4.9	1214	9	US-09-783-320-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-10-016-768-8  
Sequence 8, Application US/10016768  
Publication No. US2002014243A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT FILING DATE: 2001-10-23  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 442  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-016-768-8

Query Match 100.0%; Score 2250; DB 14; Length 442;  
Best Local Similarity 100.0%; Pred. No. 7.4e-190; Indels 0; Gaps 0;  
Matches 442; Conservative 0; Mismatches 0;

QY	1	MKKMIROFAIEYISKSGKTQENRNGSIGPSIVCKSIQMNQENSLQEOEGPLDVTNRM	60
DB	1	MKKMIROFAIEYISKSGKTQENRNGSIGPSIVCKSIQMNQENSLQEOEGPLDVTNRM	60
QY	61	QONTQOGDGVLDLSTKTKTSKSESSICDPSSNSVAGRHRREDYVERSAFADCLL	120
DB	61	QONTQOGDGVLDLSTKTKTSKSESSICDPSSNSVAGRHRREDYVERSAFADCLL	120
QY	121	SKALKDIQSGALDINKAGILYGIPOKTLLEALPGKPAFPKXKTRDFDSDYSKDSK	180
DB	121	SKALKDIQSGALDINKAGILYGIPOKTLLEALPGKPAFPKXKTRDFDSDYSKDSK	180
QY	181	ETCAVLQKVALMARQAERTKSKLNTLETSEIKFPPIASTYLYHQLTLQKWTQFKENES	240
DB	181	ETCAVLQKVALMARQAERTKSKLNTLETSEIKFPPIASTYLYHQLTLQKWTQFKENES	240

```

QY 241 LQYETSNPTVOLKIPOLRVSSSVSKSOPDGSGLDVMYOVSTSSVLBSALQKLNILPK 300
DB 241 LQYETSNPTVOLKIPOLRVSSSVSKSOPDGSGLDVMYOVSTSSVLBSALQKLNILPK 300
QY 301 QNKIECGPVTSHSSVDYFLHGDLSPLCLNSKNGTVDGTSNTEGDLDRKDSKOPRRKRG 360
DB 301 QNKIECGPVTSHSSVDYFLHGDLSPLCLNSKNGTVDGTSNTEGDLDRKDSKOPRRKRG 360
QY 361 RYRGYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKERSGTLTKTPPKKXLRP 420
DB 361 RYRGYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKERSGTLTKTPPKKXLRP 420
QY 421 DTGLYNNMTDSCGSCCKNSKRPV 442
DB 421 DTGLYNNMTDSCGSCCKNSKRPV 442

RESULT 2
US-10-029-386-33895
; Sequence 33895, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33895
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005768.16
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.85
; OTHER INFORMATION: SWISSPROT HIT: Q9YID8, EVALUATE 1.60e+00
US-10-029-386-33895

Query Match
Best Local Similarity 100.0%; Score 1411; DB 12; Length 277;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 RLHNRREDYVERSAEFADGLSLKALKDIOGALDINKAGILYGIPOKTLHLHLBALPAGK 159
DB 1 RLHNRREDYVERSAEFADGLSLKALKDIOGALDINKAGILYGIPOKTLHLHLBALPAGK 60
QY 160 PASFRNKTRDPHDSYSYKDSKETCAVLQKVALMARAQEPTEKSLNLTSEIKFPFAS 219
DB 61 PASFRNKTRDPHDSYSYKDSKETCAVLQKVALMARAQEPTEKSLNLTSEIKFPFAS 120
QY 220 TYLHQLTLQKAVTOFKENESLQYETSNPTVOLKIPOLRVSSSVSKSOPDGSGLDVMYOV 279
DB 121 TYLHQLTLQKAVTOFKENESLQYETSNPTVOLKIPOLRVSSSVSKSOPDGSGLDVMYOV 180
QY 280 SKTSSVLEGSALQKLNILPKQNKIECGPVTSHSSVDYFLHGDLSPLCLNSKNGTVDGT 339
DB 181 SKTSSVLEGSALQKLNILPKQNKIECGPVTSHSSVDYFLHGDLSPLCLNSKNGTVDGT 240
QY 340 SENTEDGDRKDSKOPRRKRGYROYDHEIMEEALAM 376
DB 241 SENTEDGDRKDSKOPRRKRGYROYDHEIMEEALAM 277

RESULT 3
US-10-016-768-2
; Sequence 2, Application US/10016768
; Publication No. US20020142443A1
; GENERAL INFORMATION:
; APPLICANT: Baehreke, Eric H.

```

```

; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
; FILE REFERENCE: 4115-131
; CURRENT APPLICATION NUMBER: US/10/016,768
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(54)
; OTHER INFORMATION: X CAN BE ANY AMINO ACID
US-10-016-768-2

Query Match
Best Local Similarity 100.0%; Score 273; DB 14; Length 53;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 353 KOPRRKRGYROYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKER 405
DB 1 KOPRRKRGYROYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKER 53

RESULT 4
US-10-016-768-3
; Sequence 3, Application US/10016768
; Publication No. US20020142443A1
; GENERAL INFORMATION:
; APPLICANT: Baehreke, Eric H.
; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
; FILE REFERENCE: 4115-131
; CURRENT APPLICATION NUMBER: US/10/016,768
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 54
; TYPE: PRT
; ORGANISM: T. nigroviridis
US-10-016-768-3

Query Match
Best Local Similarity 10.4%; Score 233.5; DB 14; Length 54;
Matches 44; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 353 KOPRRKRGYROYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKER 405
DB 1 KOPRRKRGYROYDHEIMEEALAMVMSGKMSVSKAGIYGVPHTLEYKVKER 54

RESULT 5
US-10-016-768-4
; Sequence 4, Application US/10016768
; Publication No. US20020142443A1
; GENERAL INFORMATION:
; APPLICANT: Baehreke, Eric H.
; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH
; FILE REFERENCE: 4115-131
; CURRENT APPLICATION NUMBER: US/10/016,768
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 53
; TYPE: PRT
; ORGANISM: M. musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(54)
; OTHER INFORMATION: X can be any amino acid
US-10-016-768-4

```

Query Match 10.2%; Score 229; DB 14; Length 53;  
 Best Local Similarity 81.1%; Pred. No. 5.2e-13;  
 Matches 43; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Oy 353 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAQIGYVPHSTLEYKVKER 405  
 Db 1 KHPRKGRYQYDHEIMEAIAVMGSKMSVSKAQIGYVPHSTLEYKVKER 53

RESULT 6  
 US-10-016-768-10  
 ; Sequence 10, Application US/10016768  
 ; Publication No. US20020142443A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baehrecke, Eric H.  
 ; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
 ; FILE REFERENCE: 4115-131  
 ; CURRENT APPLICATION NUMBER: US/10/016,768  
 ; CURRENT FILING DATE: 2001-10-29  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 10  
 ; LENGTH: 1165  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-016-768-10

Query Match 8.9%; Score 200.5; DB 14; Length 1165;  
 Best Local Similarity 21.7%; Pred. No. 1.9e-08;  
 Matches 92; Conservative 66; Mismatches 146; Indels 119; Gaps 14;

Oy 19 TOENRNGSIGPSIVKSIQNMQAENSLOEBOGPIDLTVNRMQEOGTQOGVLDL--ST 76  
 Db 514 SOENSNAGASLLLOQOQHQQHQQHQQOQOQOQVAAVRRHLPKSEPTETVSSLDNDAS 573  
 Oy 77 KTSIKSEESSICDPSSSENSVAGRLHRRNEDYVERSAEFADGLSLKALDKDIQSGALDINK 136  
 Db 574 EDPIKISPFKSGPSSSS-----LSP 596  
 Oy 137 AGILYGIQKTLHLHLPLPKPKSPKTRDFHDSYSYKD---SKETCAVLQKALW 192  
 Db 597 GGLVGG-----HHHPLNNSNLSISNNSN--HSSNSHRGNSRSPHSASPMLAAV- 645  
 Oy 193 ARAQARTKSKUNLLETSEIKFTASTYVLTQLTQKAVTQFKNEBLSQVETSPYQL 252  
 Db 646 --AOCGYAGNSLLTSSSSSIQKMASNIQRQI-----NEOSGQES----- 684  
 Oy 253 KIPOLRVSSVSKQOPDGGSLDVMYQ-----VSKTSVLEGSALQKLNILPKONKIECS 307  
 Db 685 ----LRNGVADDCSNNGSSSLGKPKPSIVAKIIGTDTISRFASPNLSQCH----- 735  
 Oy 308 GPVTHSSVDYFLHGDLSPLCLNSKNGTVDGTSENTEDGLDRKDS--KOPRKKGRYQY 365  
 Db 736 ----HS--AHHL-----THQOQOQOLSQALGKGRPKRGKYRNY 770  
 Oy 366 DHEIMEAIAVMGSKMSVSKAQIGYVPHSTLEYKVKERSGTLTKPPKKTLRLPDITLY 425  
 Db 771 DRDSIVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER--LMRPRKKEPKQPDV 827  
 Oy 426 NMT 428  
 Db 828 GLT 830

RESULT 7  
 US-10-016-768-1  
 ; Sequence 1, Application US/10016768  
 ; Publication No. US20020142443A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baehrecke, Eric H.  
 ; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
 ; FILE REFERENCE: 4115-131

CURRENT APPLICATION NUMBER: US/10/016,768  
 CURRENT FILING DATE: 2001-10-29  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1  
 LENGTH: 53  
 TYPE: PRT  
 ORGANISM: Drosophila melanogaster  
 FEATURE:  
 NAME/KEY: MISC FEATURE  
 LOCATION: (1)..(54)  
 OTHER INFORMATION: X can be any amino acid  
 US-10-016-768-1

Query Match 7.3%; Score 165; DB 14; Length 53;  
 Best Local Similarity 60.4%; Pred. No. 2.3e-07;  
 Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Oy 353 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAQIGYVPHSTLEYKVKER 405  
 Db 1 KGRPKGRKYNVDRDSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53

RESULT 8  
 US-10-016-768-5  
 ; Sequence 5, Application US/10016768  
 ; Publication No. US20020142443A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baehrecke, Eric H.  
 ; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
 ; FILE REFERENCE: 4115-131  
 ; CURRENT APPLICATION NUMBER: US/10/016,768  
 ; CURRENT FILING DATE: 2001-10-29  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 53  
 ; TYPE: PRT  
 ; ORGANISM: Caenorhabditis elegans  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)..(54)  
 ; OTHER INFORMATION: X CAN BE ANY AMINO ACID  
 US-10-016-768-5

Query Match 7.2%; Score 163; DB 14; Length 53;  
 Best Local Similarity 56.6%; Pred. No. 3.4e-07;  
 Matches 30; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Oy 353 KOPRKKGRYQYDHEIMEAIAVMGSKMSVSKAQIGYVPHSTLEYKVKER 405  
 Db 1 KSRPKGQYRKXKXNALDEAVRSVRGEMTVHRAGSFFGVPHSTLEYKVKER 53

RESULT 9  
 US-10-011-588-45  
 ; Sequence 45, Application US/10011588  
 ; Publication No. US20020168727A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jensen, Melody  
 ; TITLE OF INVENTION: RECOMBINANT LIGHT CHAINS OF BOTULINUM  
 ; TITLE OF INVENTION: NEUROTOXINS AND LIGHT CHAIN FUSION PROTEINS FOR USE IN  
 ; FILE REFERENCE: A34796 067252.0113  
 ; CURRENT APPLICATION NUMBER: US/10/011,588  
 ; CURRENT FILING DATE: 2002-03-29  
 ; PRIOR APPLICATION NUMBER: 09/910,186  
 ; PRIOR FILING DATE: 2001-07-20  
 ; PRIOR APPLICATION NUMBER: 09/611,419  
 ; PRIOR FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: 60/246,744  
 ; PRIOR FILING DATE: 2000-11-06

PRIOR APPLICATION NUMBER: 60/311,966  
PRIOR FILING DATE: 2001-08-09  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: FASTSEQ For Windows Version 4.0  
SEQ ID NO 45  
LENGTH: 848  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Recombinant protein encoded by SEQ ID NO:44  
US-10-011-588-45

Query Match 5.9%; Score 132.5; DB 14; Length 848;  
Best Local Similarity 20.4%; Pred. No. 0.012;  
Matches 79; Conservative 64; Mismatches 129; Indels 115; Gaps 16;

QY 25 GSIPSVYKSIQNMNAENSLQEOEGPLDITVNRMOQONTQGGDGLDSTKTSIKSE 84  
DB 273 GGHDPVSVPSTDMYVKALQNFOD-----IANRLNIVSSAQSGGI-DISLYKQIYKXK 326  
QY 85 ESSICDPSSSENSVAGRLHNRNEDYVERSAEFADGLSLKALDIOGALDINKAGILYGP 144  
DB 327 YDFEPDPKGYSV-----DKDKF-----DLTYKALMFGFTETNLAG-EYGI- 366  
QY 145 QXTLLHL-EALP-----AGKPASFKNKTRDPH-----DSYYSKDSKE 181  
DB 367 -KTRYSYSEYLPKTKTEKLDNTITYTONEGFNIAKSLKTEFNGQKAVNKEAYEESL 425  
QY 182 TCVAIQKVALMARAQERTSKNLLETSEIKFPTASTYLHQLTQKAVTQFEKESNL 241  
DB 426 EHLVIYRIAMCKPVMYKNTGKSEGCIIYNNEELFIAN-----KQSFCKDLAKARTI 477  
QY 242 QYETSNPIVQ-----LQIP-QLVSSVSKSQP 267  
DB 478 AYNTQNTIENNFSIDQLINDLSGGIDLPNENTPEPTNPDIDIPYIKQSAALKKIFV 537  
QY 268 DSGGLDVMYOVSKTSVLEGSALQKLNILPKQHK-----IECGPVTSSVDSY 318  
DB 538 DGDSEFVILHAQTFPSNI-ENQLTNSLNDALRNKKVTFSTNLVEQANTVVGAS----- 592  
QY 319 FLHGDLSPLCLNSKNGTVDG-TSENTE 344  
DB 593 -----LFTVMWVGVIDDFTSESTQ 611

RESULT 10  
US-10-029-386-32827  
Sequence 32827, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rant, David R.  
APPLICANT: Hantzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 32827  
LENGTH: 870  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO 284487.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.74  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8  
OTHER INFORMATION: SWISSPROT HIT: P46100, EVALUATE 0.00e+00  
US-10-029-386-32827

Query Match 5.8%; Score 131.5; DB 12; Length 870;  
Best Local Similarity 23.0%; Pred. No. 0.015;  
Matches 100; Conservative 63; Mismatches 173; Indels 99; Gaps 21;

QY 14 SKSGKTQENRNG-SIGPSIVKSIQNMNAENSLQEOEGPLDITVNRMOQONTQGGDGL 72  
DB 409 STSGDFDTKKGAKAKSIISKRRQTQSSSS---NYDSELEKIKSMXIGAR----- 460  
QY 73 DLSTKK--TSIKSESSICDPSSSENSVAGRLHNRNEDYVERSAEFADGLSLKALDIOG 130  
DB 461 -TTKKRIPIPTKQFDSSEDEKSKGMDNQGHKVLKTSQEGSDDAERKQERETFSAG 518  
QY 131 ALDINKAGILYGPQKLLHL-EALPAGPASEFKNTRDPHDSYSYKDSKETCAVLQKV 189  
DB 519 TVD-----KOTIHELRLRPLPKQQAAS---ASTGQVLDLSKEGQSFSLERKV 564  
QY 190 ALMARAQERTSKNLLETSEIKFPTASTYLHQLTQKAVTQFEKESNLQYETSNPT 249  
DB 565 -----AETFEKSK-----HLKTKCKKV-QDGLSDIAEKFLLKQDS--DETSDD 606  
QY 250 VOLKIPQLRVSSVSKSQPDGSGLDVMYOVSKTSVLEGSALQKLNILPKQNIKESGP 309  
DB 607 KK-----QSKKGTEERKKPS-----DFKKVYIKMEQYESSSDGTEK--LPEREET-CHFP 654  
QY 310 VTHSSVDSYFLHGDLSPLCLNSKNGTVDGTSSENTEDGLDKRDSKOPRKKRGARYQYDHEI 369  
DB 655 KGIKQI-----KNGITDG-----EKSKKIDTKSKSKKDELSDY 688  
QY 370 MEEAIAWMSGKMSVSK--AQGIYVPHSTLEVKYKERSGTLKTPPKK--LRLPDTGLY 425  
DB 689 AEKSTGAGDSDSEDEKSKNGAYG-----REKRCKLIGKSSRKQDCSSSDTEKY 740  
QY 426 NMTDSGTGSCNNSK 440  
DB 741 SMKEDG---CNSSDK 752

RESULT 11  
US-09-893-519A-37  
Sequence 37, Application US/09893519A  
Publication No. US20030027243A1  
GENERAL INFORMATION:  
APPLICANT: ANADYS PHARMACEUTICALS, INC.  
APPLICANT: THOMPSON, Craig  
APPLICANT: MOORE, Jeffrey  
APPLICANT: BUURMAN, Ed T.  
APPLICANT: BRADLEY, John  
APPLICANT: DESILVA, Thamara  
APPLICANT: HARRIS, Sandra  
APPLICANT: KOMARNITSKY, Svetlana  
APPLICANT: MENDILLO, Marc  
APPLICANT: MOORE, Daniel  
APPLICANT: MCCOY, Melissa  
APPLICANT: SANDERSON, Karen  
APPLICANT: HAO, Tariq  
APPLICANT: ZHU, Shuhao  
APPLICANT: LONG, Fan  
APPLICANT: DAVIDOV, Eugene  
TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE  
FILE REFERENCE: 0342/1G548-US2  
CURRENT APPLICATION NUMBER: US/09/893,519A  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/215,164  
PRIOR FILING DATE: 2000-06-29  
PRIOR APPLICATION NUMBER: US 60/224,457  
PRIOR FILING DATE: 2000-08-10  
NUMBER OF SEQ ID NOS: 146  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 37  
LENGTH: 534  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:

NAME/KEY: misc.feature  
OTHER INFORMATION: Corresponds to SEQ ID NO: 110  
US-09-893-519A-37

Query Match 5.6%; Score 127; DB 11; Length 534;  
Best Local Similarity 21.4%; Pred. No. 0.018;  
Matches 92; Conservative 61; Mismatches 157; Indels 120; Gaps 16;

QY 5 IROFAIEYISKSGTOENRNGSI-----GPSYCKSIQNMQAENSQEOEGPDLTLV 57  
DB 9 ISDIKIPKNDPDIJEDENASLFGHNEKNGES-----DLSYDGNSTEEYKKAHYLEV 62  
QY 58 NRMQONTQOGDGVLDLSTFK-TSISKSESSICDPSSENSVAGRLHRRNEDYVERSAEF- 115  
DB 63 -----EKSRLRAEKGLNDPKYTGKSGRQALYEEVSENEDEEEEEEKEEDALSTR 118  
QY 116 -----ADG-----LISKAL-----KDIOGALDINKAGI 139  
DB 119 TDSEDEVEIEDEESDADGETEBAQOKRHAKLSKUIQETKQAINKLSQSVQDASKG-- 176  
QY 140 LYGIPOKTL-----LHLEALPAKGPASFKNKTRDPHDSYSYKDSKETCAVLQKVALMA 193  
DB 177 -YSIIQOKTLFNNIIDRLKLOKAVIAANKPLTTESWEAKMDSEETKRLK----- 229  
QY 194 RAQAERTKSKLNLLETSEIKF-----PTASTYLHQLTLQKMYTOFEKKNESLOYETS 246  
DB 230 --ENEKLFNNLBNRLINFRIKFQLDHITQNEEVAKHKLKSKRSLELYQETNSLDSEIK 287  
QY 247 N-PTVQKLPOLRVSSVSKSQPDGGL-----LDVMYVSKTSSVLEGSALQKLN 296  
DB 288 EYRTVLNKKMSTKVSASGNALSSNKFKAINLPADVQYENQLSDMSRLMKRTKLN-RN 346  
QY 297 ILPKNKIECS-----GPVTHSSVDSYFLHGDLSPLCLNSKXGVTGTSSENEDEGID 348  
DB 347 ITPLYFOKDCANGRLPELISPVVKDSVDD-----NENSDGID 384  
QY 349 RKDSKQPRKK 358  
DB 385 IPKVDPRRK 394

RESULT 12  
US-09-924-154-16  
Sequence 16, Application US/09924154  
Patent No. US20020127241A1  
GENERAL INFORMATION:  
APPLICANT: Natum, David L.  
APPLICANT: Sim, Kim L.  
TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use  
FILE REFERENCE: 05213-0465 43170-262105  
CURRENT APPLICATION NUMBER: US/09/924,154  
CURRENT FILING DATE: 2001-08-07  
PRIOR APPLICATION NUMBER: US 60/223,525  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 972  
TYPE: PRT  
ORGANISM: Mammalian  
US-09-924-154-16

Query Match 5.5%; Score 124.5; DB 10; Length 972;  
Best Local Similarity 21.8%; Pred. No. 0.074;  
Matches 101; Conservative 74; Mismatches 187; Indels 101; Gaps 21;  
QY 32 VCKSIQMAQENSLOEBOEGPDLTVNRM---QOONTQOGDGVLDLSTKTSI----- 81  
DB 437 ICKSTVKKPDPEDIDDEFNEBSLVNPNLSLTQSQVTERVSSVDVLSIKENVLDKPRKP 496  
QY 82 -----KSEBSICDP-SENSVAGRLHRRNEDYVERSAEFADGLLSKALKDIOGALDIN 135  
DB 497 KGGTOSHVQDVGNDPSESSEKPSGA--NGREDPTSESTYNDGVITSSSLSSSGRDVS 554

QY 136 KAGILYGIPOKTLHLLEALPAKGPASFKNKTRDPHDSYSYKDSKETCAVLQKVALMARA 195  
DB 555 SSPVGVGEHEA-----KELLPPQKIIDVYTOSEDTLSQHGKSESQEOHNLDGSSL-SRH 609  
QY 196 QAERTKSKLNLLETSEIKFPPTASTYLHQLTLQKMYTOFEK-KNESLOYETSNPVQV- 252  
DB 610 SNQDEERS-----IISDVHEGHTNSLFGSQIQOQETIILSESFLTTSPEHERSKMDTHAG 665  
QY 253 --KITQLAVSSYSKQP---DGSGLL-----DVMYQVSKTSSVLEGSALQKLNILP 299  
DB 666 GKMEQVHNASVSDSSSENSNGRGGLKTKREMKGEVTVTITSKNDINLEDSYVHS----- 719  
QY 300 KONKIEGSG-----PYTHSSVDSYFLHGDLSPLCLNS---KXGT 335  
DB 720 KONKLENSGDNTQCKEHNIVLQGMDKHLENPTSERGDS-VLESEFSKLNRTSHTHDNR 778  
QY 336 VDGTSENTEDGI-----DRKDSKQPR---KKGRYRQYDHEIMEAIAVMYSGKMSYSK 386  
DB 779 IETTENNIGGLSNGSVHVDGRDSQRNRMHINSRSHGSLSDI-----VVRGD-DISN 830  
QY 387 AOGIYGVHSTLEYKVKERSGTLKTPPKKULRLPPTGLYNTMD 429  
DB 831 IEG-----GEEBEDANTLKY-PRNVLNKXNSRTYNIIE 863

RESULT 13  
US-09-882-227-624  
Sequence 624, Application US/09882227  
Publication No. US20030158396A1  
GENERAL INFORMATION:  
APPLICANT: Kleantous, Harold  
APPLICANT: Al-Garawi, Amal  
APPLICANT: Miller, Charles  
APPLICANT: Tomb, Jean-Francois  
TITLE OF INVENTION: Identification of Polynucleotides  
TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the  
FILE REFERENCE: 06132/047002  
CURRENT APPLICATION NUMBER: US/09/882,227  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 08/902,615  
PRIOR FILING DATE: 1997-07-29  
NUMBER OF SEQ ID NOS: 638  
SOFTWARE: fastSeq for Windows Version 4.0  
SEQ ID NO 624  
LENGTH: 1743  
TYPE: PRT  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 876  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-882-227-624

Query Match 5.4%; Score 121.5; DB 12; Length 1743;  
Best Local Similarity 19.8%; Pred. No. 0.33;  
Matches 98; Conservative 78; Mismatches 179; Indels 141; Gaps 18;  
QY 14 SKSGTOENRNGSGPSYCKSIQW-----NQAENSLQEOEGPDLTVNRMQOONTQ 66  
DB 202 SEGNETSSNGSLDKLFKYARKLVNKKPFTQOKNLDEETQ---LNEEDDENNEY 257  
QY 67 QGDGVLDLSTKTSIKSESSICDPSSENSVAGRLHRRNEDYVERSAEFA----- 116  
DB 258 QEETQDTLIDDETISKTKTQGHSPOLDSNEATEA--NHFENLKSSEKSSDHLNDPTET 314  
QY 117 -----DGLSKALKDIOGALDINKAGILYGIPOKTL-----LHLEAL 155  
DB 315 QTNFPGDGSSEBITD-----DSNDOELIKGSKKYITIGIIVAVLVIIILFSRSIFHYFM 368  
QY 156 PAKGPASFKNKTRDPHDSYSYKDSKETCAVLQKVALMARAQAERTKSKL----- 205

Db 369 PLEDSKSFSSKDRNLVNDIQRGE-----YRLKERNEKGNMIDKNLFNDD 418  
Qy 206 -----NLETSSEI--KEPTASTY-----LHQLTLQKAVTQFKENESLQYET 245  
Db 419 PNRLTYLNTAINEAIEDKNPLRAFYECISNGAYEBCLTKLIDKLODMKKTLEAYNCI 478  
Qy 246 SNPTVQLTIQOLRVASVSKSDPGSLDVMYQVSKTSVLEGS-----ALQKLNILPKQ 301  
Db 479 KN-----AKTEERIKCIDLKIDENLKKSLNQKQVQVALDCLKNAKTDE 523  
Qy 302 NKIECSGPTVTHSSVSYF-----LHGDISPL--CLNS-----KNGTVDGTSENTEDGID 348  
Db 524 ERNECTKLINPDIHEKFRKELELQKELQYKDCIKMKTAEKVKCKLGSLKEAIERLK 583  
Qy 349 R-----KSKOPKRGGRYROYDHEIMEEALA--MWSGKMSVSKAGIYGVPHSTLEY 400  
Db 584 QOALDCLNANKATDERNECLKNIPDQLQKELLADMSVAKYKDCVSKAR-----NE 633  
Qy 401 KVKERSGTLTKTPPKKK 416  
Db 634 KEKOCEKULTPEARK 649

RESULT 14  
US-10-309-933-4  
; Sequence 4, Application US/10309933  
; Publication No. US20030162203A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsumoto, Naomichi  
; APPLICANT: Niihawa, No. US20030162203A1  
; TITLE OF INVENTION: NUCLEIC ACID, PROBE COMPRISING THE NUCLEIC ACID AND SCREENING MET  
; FILE OF INVENTION: USING THE PROBE  
; FILE REFERENCE: 782 229  
; CURRENT APPLICATION NUMBER: US/10/309,933  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2696  
; TYPE: PRT  
; ORGANISM: human chromosome  
US-10-309-933-4

Query Match 5.3%; Score 120; DB 12; Length 2696;  
Best Local Similarity 22.0%; Pred. No. 0.88;  
Matches 108; Conservative 61; Mismatches 22; Indels 100; Gaps 20;

Qy 1 MKMIRQPAIEYISKGTQENRNGSIGPISYCKSIQMOAENSLOEQGPLDLTVRM 60  
Db 511 VKKGHIQFEAHKDERGKIPE-----LGLNFIGDISPTQASNELSR-----IANSI 558  
Qy 61 QEOUQOQDGVLDLSTKTSIKSESSICDPSSENSVAGRLHNRNEDYVERSAEPADGL 120  
Db 559 TGSNAPSPSFFSSCGKNTAKKEFETSGND----- 588  
Qy 121 SKALKDIOGSLDILKAGILVIGIPKTL-----LHLEALPACKPASFKNKTRDFHDS 173  
Db 589 --SLIGPEGAL--ISKCREKNKPORSIKVCSKYKLCYIGADEBEKSDSISICTSDG 645  
Qy 174 VSYKDSKETCAVLQKVALMARQAERTESKLNLETSEIKPT--ASTYLLHQLTLQKAVT 232  
Db 646 SSDLDPIHSSSDNSVLEIPDAFPRTE--NMLSMQKNKIKYSRAAINTVAKAKQKLI 704  
Qy 233 QPKENESL-----QYETSNPTVQLKI POLRVSS--VSKSQPDG-----SGLLDVMYQVSKT 282  
Db 705 SNSHTDLMGCTKSAEPGETESQVNLSDLKASTLVHKQSDPTNDALSPKFNLSISSISE 764  
Qy 283 SSVLEG-----SALQKLNILPKQNKIEC-----SGPYTHSSVDVSYFLHGLD--SPL 327  
Db 765 NSLIGGANQALLHSHSKQKOPFRSIRKIKHENPYMAEPVAINESCSLCCSSDTPKGSPL 824  
Qy 328 CLNSKNGTVDGTS-----ENTEDGLDRKDS-----KOPRKRGRGRYQYDHEIMEEAIMV 377

Db 825 ASISKSGVDSGLKLLNMHEKTRSDSIETAVVGHVLSLEKLSYRSLGEVDSGTSK- 883  
Qy 378 MSGKMSVSKAGIYGV-----HSTLEYKVKESGTLKTPPKKALPDTGLYNNMTDSG 431  
Db 884 PSKELFSSASSQNHIEIPDPYKFTLLMMKDMHDS--KT--KEORLMTAQNLVSYRSPG 940  
Qy 432 TGSCKNSKPV 442  
Db 941 RGDCTNS-PV 950

RESULT 15  
US-10-032-585-7319  
; Sequence 7319, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7319  
; LENGTH: 1240  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-10-032-585-7319

Query Match 5.3%; Score 119.5; DB 12; Length 1240;  
Best Local Similarity 22.0%; Pred. No. 0.3;  
Matches 93; Conservative 62; Mismatches 155; Indels 113; Gaps 18;

Qy 37 QMOAENSLOEQGPLDLTVRMQOQDGVLDLSTKTSIKSESSICDPSSENS 96  
Db 698 ELNOKIRKLOPKR--PKDLEINLAEETIGALDLDKLPVLRNOKTSIE----- 741  
Qy 97 VAGLHNRNEDYVERSAEPADGLSKALKDIOGALDINKAGILYGIPOKTLHLHALP 156  
Db 742 --RIKDRSEI-----EFQGLFKGFDKSIQKQNEITK--INGKIDKV--NEMMK 787  
Qy 157 AGKPAEKNTKRDHDSYKDSKETCAVLQKVALMARQAERTESKL----- 205  
Db 788 SSKDLIF--AEFCERYGFVNGIEDYENMHGSLTRVRAR--ERAQFSKTSIVLONKLD 842  
Qy 206 --NLETSSEIKFPASTYLLH--QLTLQKAVTQFKENESL-----QYETSNPTV----- 250  
Db 843 KERLETKDRKRSIESLIVLEDDLAKVLTETKGLUESLDAKAEYEVLOTETIQFPDSM 902  
Qy 251 --QKIPOLRVASVSKSQPDGSLDVMYQVSKTSVLEGSALQKLNILPKQNKIECSG 308  
Db 903 QSQLKTSKIESDLDKSLVSTLVKEITQLEENLTKTDSERAVLHNC-----KIQ-- 954  
Qy 309 PVTSSVSDVSYFLHDLPLCLNSKNGTVDGTSENTEDGLRKDSKOPRKRGRYQYDHE 368  
Db 955 --NINPLIDDLDSI-----SVGENVLESSI-----KEVYKIEIDYE 989  
Qy 369 IMEBAIMWMSGKMSVSKAGIYGVPHSTLEYKVKESGTLK--TPPK--KLRLPDTG 423  
Db 990 MLEERFKEVFNKL-----OSELEVLQNTISDLEKLTPAKAIERLREYETK 1037  
Qy 424 LYN 426  
Db 1038 LRN 1040

Search completed: October 28, 2003, 12:17:02  
Job time : 84.1495 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 56.2545 Seconds  
(without alignments)  
2027.556 Million cell updates/sec

Title: US-10-016-768a-8

Perfect score: 2250

Sequence: 1 MKKMIROFAIEYISKSGKTQ.....GLYNTDSGTSCSKSKSPV 442

Scoring table:

BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PLANT:\*  
10: SP\_PLANT:\*  
11: SP\_ROTENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2114.5	94.0	517	11	Q8CJG4 mus musculu
2	1090	48.4	213	4	Q96NKK1 homo sapien
3	1089	48.4	393	11	Q8C9J6 mus musculu
4	502	22.3	320	4	Q8N3X6 mus musculu
5	501.5	22.3	433	11	Q8BGT2 mus musculu
6	499.5	22.2	572	4	Q96JNO mus musculu
7	497.5	22.1	619	4	Q8N3L6 mus musculu
8	470	20.9	396	11	Q8C9Q0 mus musculu
9	444	19.7	223	11	Q8C9B1 mus musculu
10	444	19.7	315	11	Q8BRT8 mus musculu
11	444	19.7	315	11	Q8BRT7 mus musculu
12	200.5	8.9	1165	5	Q9VDE0 drosophila
13	193.5	8.6	1598	5	Q95YH8 apis mellif
14	183	8.1	185	5	Q22051 caenorhabdi
15	141	6.3	689	10	Q9FNZ7 oryza sativ
16	133.5	5.9	1109	6	Q00756 oryctolagus

17	133.5	5.9	3616	13	Q9M6V0 gallus galli
18	130.5	5.8	678	5	O61493 drosophila
19	129.5	5.8	545	5	O17584 caenorhabdi
20	129.5	5.8	1591	11	P97868 mus musculu
21	129.5	5.8	2152	6	Q8MJ06 papio hamad
22	129.5	5.8	3099	5	Q8MYH0 dictyosteli
23	128	5.7	1256	5	Q22126 caenorhabdi
24	127	5.6	534	3	Q06631 saccharomyc
25	126.5	5.6	983	12	Q69530 human herpe
26	126.5	5.6	1078	12	Q9QJ15 human herpe
27	126.5	5.6	4493	5	Q8WPA9 dictyosteli
28	126	5.6	673	3	Q9P7X7 schizosach
29	126	5.6	430	5	Q8WQZ7 calliphora
30	125.5	5.6	1819	16	Q9ZLV0 helicobacte
31	125.5	5.6	2308	5	Q9VP17 drosophila
32	125	5.6	948	3	Q94603 schizosach
33	124.5	5.5	631	5	O81K15 plasmodium
34	124.5	5.5	983	12	Q69532 human herpe
35	124.5	5.5	983	12	Q69531 human herpe
36	124.5	5.5	1388	5	O81FM3 plasmodium
37	124.5	5.5	1444	5	Q9VTN2 drosophila
38	124.5	5.5	1514	5	Q88Y55 drosophila
39	124.5	5.5	2954	13	O42263 xenopus lae
40	124.5	5.5	3187	11	O63714 ratius norv
41	124	5.5	18519	5	O81SF6 caenorhabdi
42	124	5.5	18534	5	O81SF7 caenorhabdi
43	123.5	5.5	787	5	Q8MYH8 dictyosteli
44	123.5	5.5	1671	5	Q8WQ60 caenorhabdi
45	123.5	5.5	1827	5	Q20042 caenorhabdi

#### ALIGNMENTS

RESULT 1	
Q8CJG4	PRELIMINARY; PRT; 517 AA.
AC Q8CJG4;	
DT 01-MAR-2003 (TREMBLrel. 23, Created)	
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)	
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE Transcription factor MLR1.	
GN MLR1.	
OS Mus musculus (Mouse)	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX NCBI_TaxID=10090;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=Brain;	
RT TISSUE=Brain;	
RA Kuneda T., Park J., Takeuchi H., Kubo T.;	
RT "Mus musculus mlr1 and mlr2 mRNA for transcription factor MLR1 and	
RT MLR2."	
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AB076078; BAC20954.1; -	
SO SEQUENCE 517 AA; 57316 MW; C97403D3D296C52E CRC64;	

Query Match	94.0%; Score 2114.5; DB 11; Length 517;
Best Local Similarity	94.1%; Pred. No. 2e-129;
Matches	416; Conservative 12; Mismatches 13; Indels 1; Gaps 1;
QY	1 MKKMIROFAIEYISKSGKTQENRNGSIQPSIVCKSIQNMQVENSIOEQEGPLDTVTRM 60
DB	77 MKKMIROFAIEYISKSGKTQENRNGSIQPSIVCKSIQNMQVENSIOEQEGPLDTVTRT 136
QY	61 QEQNTQGDGVLDSTKTKTSIKSEESSICDPSSESVAGRLHRNEDIVERSAEADGL 120
DB	137 QEQNTQGDGVLDSTKTKTSIKSEESSISDPSSESVAGRLHRNEDIVERSAEADGL 196
QY	121 SKALDIOSGALDINKAGILYGIPOKTLHLHLALPCKPASPKNKTRDFHDSYKXSK 180
DB	197 SKALDIOSGALDINKAGILYGIPOKTLHLHLALPCKPASPKNKTRDFHDSYKXSK 256

```
OY 181 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 240
DB 257 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 316
OY 241 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 300
DB 317 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 376
OY 301 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 360
DB 377 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 436
OY 361 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 420
DB 437 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 496
OY 421 DTGLYNTDSTGSGCKNSKSPV 442
DB 497 DTGLY-NTDSTGSGCKNSKSPV 517
```

## RESULT 2

```
O96NKL PRELIMINARY; PRT; 213 AA.
AC Q96NKL;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ30696.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Tashiro H., Yamazaki M., Matanabe K., Kumagai A., Itakura S.,
RA Tashiro H., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Ohtsuki T., Sato H., Ota T., Wakamatsu A., Ichii S., Yamamoto J.,
RA Isono Y., Kawai H., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagaastma M., Matsuo K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuno Y., Nagai K., Isegai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK055258; BAB70892.1; -.
KW Hypothetical protein.
SQ SEQUENCE 213 AA; 23477 MW; 4D7F6CABF95251B2 CRC64;
```

Query Match 48.4%; Score 1090; DB 4; Length 213;  
Best Local Similarity 99.5%; Pred. No. 2,3e-63;  
Matches 212; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
OY 220 MYTQKEKNEISLQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGS 289
DB 1 MYTQKEKNEISLQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGS 60
OY 290 ALQKLNILPKQNKIEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLDR 349
DB 61 ALQKLNILPKQNKIEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLDR 120
OY 350 KDSKQPRKKRGRYROYDHEIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTL 409
DB 121 KDSKQPRKKRGRYROYDHEIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTL 180
OY 410 KTPPKKKLRLPDTGLYNTDSTGSGCKNSKSPV 442
DB 181 KTPPKKKLRLPDTGLYNTDSTGSGCKNSKSPV 213
```

## RESULT 3

```
OY 181 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 240
DB 257 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 316
OY 241 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 300
DB 317 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 376
OY 301 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 360
DB 377 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 436
OY 361 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 420
DB 437 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 496
OY 421 DTGLYNTDSTGSGCKNSKSPV 442
DB 497 DTGLY-NTDSTGSGCKNSKSPV 517
```

## RESULT 4

```
O98N3X6 PRELIMINARY; PRT; 393 AA.
AC O98N3X6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK041987; BAC31123.1; -.
KW Hypothetical protein.
FT NON TER 393
SQ SEQUENCE 393 AA; 43892 MW; 3742CF675978C6C3 CRC64;
```

Query Match 48.4%; Score 1089; DB 11; Length 393;  
Best Local Similarity 92.3%; Pred. No. 5.9e-63;  
Matches 216; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

```
OY 181 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 240
DB 257 ETCAVLQKVALMARQAERTKSKLNLTSEIKPEPTASTYLHQLTQKMTOPKEKES 316
OY 241 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 300
DB 317 LQYETSNPTVOLKIPOLRVSSVSKSQDPGSGLLDMVQVSKTSSVLEGGALQKLNILPK 376
OY 301 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 360
DB 377 QNKTEGSGPVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTEDGLRKSKOPKRRKG 436
OY 361 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 420
DB 437 RYRQVDHIMEEALAMWMSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLTKPPKKKLRLP 496
OY 421 DTGLYNTDSTGSGCKNSKSPV 442
DB 497 DTGLY-NTDSTGSGCKNSKSPV 517
```

## RESULT 5

Query Match 22.3%; Score 502; DB 4; Length 320;  
Best Local Similarity 89.5%; Pred. No. 6.2e-25;  
Matches 102; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Db 163 MKMIRQFAIYISKGTQENRNGSIQPSIVCKSIQMOAENSIOEOPDLTVNRM 222  
Qy 61 QEONTQOGDVLSTKTKTSIKSESSICDPSSSENSVAGRLHRNEDYVERSAAE 114  
Db 223 QEONTQOGDVLSTKTKTSIKSESSICDPSSSENSVAGRLHRNEDYVERSAAE 276

## RESULT 5

Q9BGT2 PRELIMINARY; PRT; 433 AA.  
ID Q9BGT2  
AC Q9BGT2 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Transcription factor MLR2 (Hypothetical protein).  
GN MLR2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kuneda T., Park J., Takeuchi H., Kubo T.,  
RT "Mus musculus mlr1 and mlr2 mRNA for transcription factor MLR1 and  
MLR2."  
RT Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AB076079; BAC20955.1;  
DR EMBL; AK041090; BAC30816.1;  
KW Hypothetical protein.  
SQ SEQUENCE 433 AA; 47124 MW; 736656D1F7E9A041 CRC64;

Query Match 22.3%; Score 501.5; DB 11; Length 433;  
Best Local Similarity 34.5%; Pred. No. 9.8e-25;  
Matches 162; Conservative 62; Mismatches 143; Indels 103; Gaps 22;

Qy 1 MKMIRQFAIYISKGTQENRNGSIQPSIVCKSIQMOAENSIOEOPDLTVNRM 48  
Db 1 MORMIOQFAIYISKGTQENRNGSIQPSIVCKSIQMOAENSIOEOPDLTVNRM 60  
Qy 49 QEPLDVLTVNMOBONTQOGDVLSTKTKT-----TSIKSESSICDPSSSENSVAGRLHRN 104  
Db 61 QDSPLDVLTVNMOBONTQOGDVLSTKTKT-----TSIKSESSICDPSSSENSVAGRLHRN 119  
Qy 105 REDYVERSAEFAADGLSKALD-----IOSGALDINKAGILYGIPOKTLHLHLALPAG 158  
Db 120 RPD-----GLRSGDGVPPRSLODGTREGFGHSTSLKVLPA-----RSLQISEELSRN 167  
Qy 159 K-----PASFGKTRDPFDSYSYKDSKETCAVLQVLMARAOAE-RTEKSKLN--- 206  
Db 168 QLSAASLGPESGLN-----HGOH-----LILREASWAKPHYEFNLSRMKFRNG 213  
Qy 207 -LLETSEIKPPTASTYHLQTLQKMTVTFKEKNESLOYETSNPTVOLKIPOLARVSVSKS 265  
Db 214 ALSNISDLPFLAENS-----APPKAHTKODGKR-DMSHSSP-VDLKIPVGRGMDLSWE 266  
Qy 266 QPDGSGLLDVMYQVSKTSVLT-----EGSALQKLNILPKONKITEC--SGPYTHSSVDSYF 319  
Db 267 SRTGD-----QYSYSLVWGSQTESALSKKLRAILPKONKSKMLDAGP-----DSWG 313  
Qy 320 LHGDLSPCLNLSKGTGTSENTEDGLDRKDSKOPKRGKRYROYDHEIMEAIIAMWS 379  
Db 314 SDAE-----OSTSGQPYPTSDQEGD-----PGSKOPRKRKRGRYROYNSEIIEEALISVMS 363

Qy 380 GKMSVSAQAGIYGVPHSTLEKYKERSGTLKTPPKKLRL-----PTGL 424  
Db 364 GKMSVSAQAGIYGVPHSTLEKYKERSGTLKTPPKKLRL-----PTGL 413

## RESULT 6

Q96JNO PRELIMINARY; PRT; 572 AA.  
ID Q96JNO  
AC Q96JNO  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Hypothetical protein KIAA1795 (Fragment).  
GN KIAA1795.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=21245130; PubMed=11347906;  
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.,  
RT "Prediction of the coding sequences of unidentified human genes. XX.  
The complete sequences of 100 new cDNA clones from brain which code  
for large proteins in vitro."  
RL DNA Res. 8:85-95(2001).  
DR EMBL; AB058698; BAB47424.1;  
KW Hypothetical protein.  
FT NON TER  
SQ SEQUENCE 572 AA; 62730 MW; FB0A401D3F060DF4 CRC64;

Query Match 22.3%; Score 499.5; DB 4; Length 572;  
Best Local Similarity 33.6%; Pred. No. 1.9e-24;  
Matches 158; Conservative 66; Mismatches 143; Indels 103; Gaps 20;

Qy 1 MKMIRQFAIYISKGTQENRNGSIQPSIVCKSIQMOAENSIOEOPDLTVNRM 48  
Db 140 MORMIOQFAIYISKGTQENRNGSIQPSIVCKSIQMOAENSIOEOPDLTVNRM 199  
Qy 49 QEPLDVLTVNMOBONTQOGDVLSTKTKT-----TSIKSESSICDPSSSENSVAGRLHRN 104  
Db 200 QDSPLDVLTVNMOBONTQOGDVLSTKTKT-----TSIKSESSICDPSSSENSVAGRLHRN 258  
Qy 105 REDYVERSAEFAADGLSKALD-----IOSGALDINKAGILYGIPOKTLHLHLALPAG 158  
Db 259 RPD-----GLRSGDGVPPRSLODGTREGFGHSTSLKVLPA-----RSLQISEELSRN 306  
Qy 159 K-----PASFGKTRDPFDSYSYKDSKETCAVLQVLMARAOAE-RTEKSKLN--- 206  
Db 307 QLSAASLGPESGLN-----HGOH-----LILREASWAKPHYEFNLSRMKFRNG 352  
Qy 207 -LLETSEIKPPTASTYHLQTLQKMTVTFKEKNESLOYETSNPTVOLKIPOLARVSVSKS 265  
Db 353 ALSNISDLPFLAENSAPPKAALQ-----AKQDGKDVSHSSPVDLKIPOVGRGMDLSWE 405  
Qy 266 QPDGSGLLDVMYQVSKTSVLT-----EGSALQKLNILPKONKITEC--SGPYTHSSVDSYF 319  
Db 406 SRTGD-----QYSYSLVWGSQTESALSKKLRAILPKONKSKMLDAGP-----DSWG 452  
Qy 320 LHGDLSPCLNLSKGTGTSENTEDGLDRKDSKOPKRGKRYROYDHEIMEAIIAMWS 379  
Db 453 SDAE-----OSTSGQPYPTSDQEGD-----PGSKOPRKRKRGRYROYNSEIIEEALISVMS 502  
Qy 380 GKMSVSAQAGIYGVPHSTLEKYKERSGTLKTPPKKLRL-----PTGL 424  
Db 503 GKMSVSAQAGIYGVPHSTLEKYKERSGTLKTPPKKLRL-----PTGL 413

## RESULT 7

Q9N3L6 PRELIMINARY; PRT; 619 AA.  
ID Q9N3L6  
AC Q9N3L6

DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFPA51A142.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mambut R., Heubner D., Mewes H.W., Well B., Wiemann S.;  
 RI Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL834245; CAD38921.1; -  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 619 AA; 67378 MW; 7912866CF8A5110 CRC64;  
 Query Match 22.1%; Score 497.5; DB 4; Length 619;  
 Best Local Similarity 33.4%; Pred. No. 2.8e-24;  
 Matches 157; Conservative 64; Mismatches 146; Indels 103; Gaps 19;  
 QY 1 MKKMIROFAIEYISKSGKTQE-----NRNGS-----IGSIYCKSIQNNQAEISIQEE 48  
 Db 187 MORMIOQFAEYTSKNSSTQDPSQPNSTKNQSLPKASPVTTSPATATQNPVLISKLMAD 246  
 QY 49 QEGPLDITVNRMOEQNTQCGGVLDLSTKKT---TSKSEESSICDPSSNSVAGRHRN 104  
 Db 247 QDSPLDLTVRKSSQSEPEEQ-DGVLDLSTKSPCAGSTLSHSPGSSSTQNGRGRPSQY 305  
 QY 105 REDYVERSAEFAADGLSKALKD-----IOSGALDINKAGILVGIPOKTLHLHLALPAG 158  
 Db 306 RPD-----GLRSGGVPRPSLQDGTREGFGHSTLKVPLA-----RSLQISEELLRN 353  
 QY 159 K-----PASFKNKTRDFHDSYSYKSKETCAVLQVAMARAQAE-RTEKSKLN--- 206  
 Db 354 QLSTAASLGSGIQN-----HGOH-----LILSREASMAKPHYEFSLMKFRNG 399  
 QY 207 -LLETSEIKPTASTYLHQLTLQKMTQPKENESLQYETSNPTVOLKIPOLRAVSVYSKS 265  
 Db 400 ALSNISDLPFLAENSAPFKMALQ-----AKQDCKKQVSHSSPVDLKIPOVRGMDLSWE 452  
 QY 266 QPDGSSGLDVMYQVSKTSVYL-----EGSALQKKNILPKONKIEC--SGPVTHSSVDSYF 319  
 Db 453 SRTGD-----QYSYSLVWGSQTESALSCKLRALILPKQGRKSMLDGP-----DSWG 499  
 QY 320 LHGDLSPCLNSKNGYDGTSENTEDGLDRKDSKQPKRKGROYDHEIMEEAIAMVMS 379  
 Db 500 SDAEQS-----TPGQVPYPTSDQGDGDPGSKQPRKRGROYNSEILFEAISVYMS 549  
 QY 380 GKMSVSKAOGIYGVPHSTLEYKVKERSGTLKTPPKKRL-----PDGGL 424  
 Db 550 GKMSVSKAOGIYGVPHSTLEYKVKERLGLTKNPPKKMKLMRSEGPDSV 599  
 RESULT 8  
 Q8C900 PRELIMINARY; PRT; 396 AA.  
 ID Q8C900;  
 AC Q8C900;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RI Nature 420:563-573(2002).  
 DR EMBL; AK041621; BAC31007.1; -  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 396 AA; 43085 MW; EE4A585F62336B35 CRC64;  
 Query Match 20.9%; Score 470; DB 11; Length 396;  
 Best Local Similarity 34.2%; Pred. No. 9.7e-23;  
 Matches 155; Conservative 59; Mismatches 141; Indels 98; Gaps 21;  
 QY 1 MKKMIROFAIEYISKSGKTQE-----NRNGS-----IGSIYCKSIQNNQAEISIQEE 48  
 Db 1 MORMIOQFAEYTSKNSSTQDPSQPNSTKNQSLPKASPVTTSPATATQNPVLISKLMAD 60  
 QY 49 QEGPLDITVNRMOEQNTQCGGVLDLSTKKT---TSKSEESSICDPSSNSVAGRHRN 104  
 Db 61 QDSPLDLTVRKSSQSEPEEQ-DGVLDLSTKSPCAGSTLSHSPGSSSTQNGRGRPSQY 119  
 QY 105 REDYVERSAEFAADGLSKALKD-----IOSGALDINKAGILVGIPOKTLHLHLALPAG 158  
 Db 120 RPD-----GLRSGGVPRPSLQDGTREGFGHSTLKVPLA-----RSLQISEELLRN 167  
 QY 159 K-----PASFKNKTRDFHDSYSYKSKETCAVLQVAMARAQAE-RTEKSKLN--- 206  
 Db 168 QLSTAASLGSGIQN-----HGOH-----LILSREASMAKPHYEFSLMKFRNG 213  
 QY 207 -LLETSEIKPTASTYLHQLTLQKMTQPKENESLQYETSNPTVOLKIPOLRAVSVYSKS 265  
 Db 214 ALSNISDLPFLAENS-----APFKMAHQTKQDKR-DMSHSSP-VDLKIPQVRGMDLSWE 266  
 QY 266 QPDGSSGLDVMYQVSKTSVYL-----EGSALQKKNILPKONKIEC--SGPVTHSSVDSYF 319  
 Db 267 SRTGD-----QYSYSLVWGSQTESALSCKLRALILPKQGRKSMLDGP-----DSWG 313  
 QY 320 LHGDLSPCLNSKNGYDGTSENTEDGLDRKDSKQPKRKGROYDHEIMEEAIAMVMS 379  
 Db 314 SDAE-----QSTSGQVPYPTSDQGDGDPGSKQPRKRGROYNSEILFEAISVYMS 363  
 QY 380 GKMSVSKAOGIYGVPHSTLEYKVKERSGTLKTP 412  
 Db 364 GKMSVSKAOGIYGVPHSTLEYKVKERLGLTKNPPKKMKLMRSEGPDSV 396  
 RESULT 9  
 Q8C9B1 PRELIMINARY; PRT; 223 AA.  
 ID Q8C9B1;  
 AC Q8C9B1;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=C57BL/6J; TISSUE=Cerebellum;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RI Nature 420:563-573(2002).  
 DR EMBL; AK042567; BAC31295.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 223 AA; 24472 MW; B019FF8BFCB7C72F CRC64;  
 Query Match 19.7%; Score 444; DB 11; Length 223;  
 Best Local Similarity 80.4%; Pred. No. 2.3e-21;  
 Matches 90; Conservative 7; Mismatches 15; Indels 0; Gaps 0;



DI VI-NOV-1990 (11)

DT	01-NOV-1996 (T-EMBLrel. 01, Created)
DT	01-NOV-1996 (T-EMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)  
 DE T01C1.3 protein.  
 GN T01C1.3  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lennard N.  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018 (1998).  
 DR EMBL; Z68010; CAA92009.1; -  
 DR WormPep; T01C1.3; CE03594.  
 SQ SEQUENCE 185 AA; 20706 MW; F9F59327B318F641 CRC64;

Query Match 8.1%; Score 183; DB 5; Length 185;  
 Best Local Similarity 32.4%; Pred. No. 0.00016;  
 Matches 47; Conservative 30; Mismatches 50; Indels 18; Gaps 5;

QY 263 SKSQDPSGLDWMYQVSKTSSVLEGSALQKL-KNLPKONKIECSGPVTHSSVDSYFLH 321  
 DB 9 TNSLEGTEPREMD-KKSCSPLDPKWLESIMQNLFKTQGNV--PDSANISNVDT 64  
 QY 322 GDLSPCLNKKNGTVDGTSNTEDGLDRKDSKOPRKRGROYDHEIMEEALIAMVSGK 381  
 DB 65 ---TTPISSEKQKHGNE-----WKSRKQKQYKYNALDEAVRSVRGE 111  
 QY 382 MSVSKAQGIYGVPHSTLEYKERS 406  
 DB 112 MTHRAGSFVGVPHSTLEYKERN 136

## RESULT 15

ID Q9FNZ7 PRELIMINARY; PRT; 689 AA.  
 AC Q9FNZ7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE P0038C05.27 protein.  
 GN P0038C05.27.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Saeki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
 RT clone:P0038C05."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003044; BAB19354.1; -  
 DR Gramene; Q9FNZ7; -  
 DR InterPro; IPR000253; FHA;  
 DR Pfam; PF00498; FHA; 1.  
 DR SMART; SM00240; FHA; 1.  
 DR PROSITE; PS50006; FHA DOMAIN; 1.  
 SQ SEQUENCE 689 AA; 75868 MW; C75474B0A9668940 CRC64;

Query Match 6.3%; Score 141; DB 10; Length 689;  
 Best Local Similarity 21.7%; Pred. No. 0.47;  
 Matches 103; Conservative 64; Mismatches 174; Indels 134; Gaps 19;

QY 1 MKMIRQFAIEYISKSGKTQ-----ENRNGSIGPSIVCKSI 36

DB 262 MKKEIDATIRADISQGGTLTQOQTOIARNEQRTSQLMELENTLETTLNDISRESIGARTG 321  
 QY 37 QMNAEN--SIQEOEGFLD-----LTNRMOEQNTQCGDVL-----DLS 75  
 DB 322 NSNRSHKASLEEEEDDLSDEDDFYDRTKKSSSHKSSQOQVETADSLDKKDTITSIE 381  
 QY 76 TKKTSIKSESSICDPSSENVAGRLHNRREDYERSAEFADGLSKALKDIOGALDIN 135  
 DB 382 SKKLVEEKKLA--KSENADV-----DLDAYMGSLSQLVHDKIAQIQKELSDIQ 433  
 QY 136 KAGILYGIPOKTLHLLEALPAGKPAFKN--KTRDFHDSYKSKETCAVLQKVALMAR 194  
 DB 434 TE-----LGRVYLLKI-ADPMGEARARDKPRETKSPASVDSLRPSRKQNV---AQ 484  
 QY 195 AOAERTESKLNLETSEIKFPPTASTYVHLQTLQKMTQFKKESLOYET-SNPVQIK 253  
 DB 485 NKASTEERLKESCAEKTQVDPKAE-----BEKGISTNOENSGKPAFSP 528  
 QY 254 IPO-----LRVSSVSKQDPDGSGLDWMYQVSKTSSVLEGS----- 290  
 DB 529 KPQWLGDKRTVSESENCIKESANEERTDNF---VDYKDKRT--ILSGSANGKDLLEEA 582  
 QY 291 ----LQKLKNLPKONKIECSGPVTHSSVDSYFLHGLDSPCLNKKNGTVDGTSNTEDG 346  
 DB 583 PGLIRKRRKSDQSAANEV-----SSVESEASAADVALLIKHKRL--QTSDEME 633  
 QY 347 LDRKDSKOPRKRGROY-----DHEIMEEALIAMVSGKMSVSKAQG 389  
 DB 634 NEPQASRRKSKSKQKRVLGPARPDLDPAGPDHETWVPPEGQTGGRISLNDRLG 688

Search completed: October 28, 2003, 12:04:26  
 Job time : 61.2545 secs

**THIS PAGE BLANK (USPTO)**



PI Baehrecke EH;  
 XX WPI; 2002-479717/51.  
 DR  
 XX  
 PT Novel programmed cell death modulating proteins, useful for treating or  
 PT preventing disorders associated with abnormal cell proliferation and  
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
 PT interaction -  
 XX  
 PS Claim 1; Fig 1; 88pp; English.  
 XX  
 CC The present invention relates to novel programmed cell death modulating  
 CC proteins and polynucleotides encoding such proteins. Sequences of the  
 CC invention are useful to screen potential cellular apoptosis inhibiting  
 CC compounds to determine their use as therapeutic agents for treatment of  
 CC diseases associated with increased programmed cell death. They are also  
 CC useful for treating or preventing disorders associated with decrease in  
 CC apoptosis. Programmed cell death modulating sequences are useful for  
 CC treating or preventing cancer e.g., adenocarcinoma, leukaemia, lymphoma,  
 CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
 CC invention are useful for treating disorders associated with increase  
 CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
 CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
 CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
 CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
 CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
 CC diseases and other infectious or genetic immunodeficiencies. Sequences  
 CC of the invention are used as vaccines and in gene therapy. The present  
 CC sequence is fruit fly E93 programmed cell death modulating protein  
 CC conserved domain.  
 CC  
 XX  
 SO Sequence 53 AA;  
 XX  
 Query Match 100.0%; Score 278; DB 22; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 2,7e-32;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KGTFRKRGKRYNDRDLSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 1 KGTFRKRGKRYNDRDLSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 XX  
 RESULT 2  
 AB71145  
 ID ABB71145 standard; Protein; 1140 AA.  
 AC ABB71145;  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 40227.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 FN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 XX WPI; 2001-656860/75.  
 DR  
 DR N-PSDB; ABL15248.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 40227; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (AB157737-AB172072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SO Sequence 1140 AA;  
 XX  
 Query Match 100.0%; Score 278; DB 22; Length 1140;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-30;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KGTFRKRGKRYNDRDLSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 741 KGTFRKRGKRYNDRDLSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 793  
 XX  
 RESULT 3  
 AAE24372  
 ID AAE24372 standard; Protein; 1165 AA.  
 AC AAE24372;  
 XX  
 XX 04-OCT-2002 (first entry)  
 DT  
 XX  
 DE Fruit fly E93 programmed cell death modulating protein #1.  
 XX  
 KW Fruit fly; programmed cell death modulating protein; adenocarcinoma;  
 KW cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
 KW neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
 KW Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
 KW aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
 KW reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
 KW vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
 KW myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective;  
 KW cardiant; cancer; E93 protein.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 FN WO200234882-A2.  
 XX  
 PD 02-MAY-2002.  
 XX  
 PF 29-OCT-2001; 2001WO-US48053.  
 XX  
 PR 27-OCT-2000; 2000US-243865P.  
 XX  
 PA (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
 XX  
 PI Baehrecke EH;  
 XX  
 XX WPI; 2002-479717/51.  
 DR  
 DR N-PSDB; AAD39237.  
 XX  
 PT Novel programmed cell death modulating proteins, useful for treating or  
 PT preventing disorders associated with abnormal cell proliferation and  
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
 PT infarction -  
 XX  
 PS Claim 9; Page 65-71; 88pp; English.  
 XX

CC The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is fruit fly E93 programmed cell death modulating protein.  
XX  
SQ Sequence 1165 AA;

Query Match 100.0%; Score 278; DB 23; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 1.4e-30;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KGTREPKRGKRYNYDRDSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
DB 758 KGTREPKRGKRYNYDRDSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 810

RESULT 4  
AAE24595  
ID AAE24595 standard; Protein; 53 AA.

AC AAE24595;

DT 04-OCT-2002 (first entry)

DE Nematode E93 programmed cell death modulating protein conserved domain.

XX -Nematode; programmed cell death modulating protein; adenocarcinoma;  
XX cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
XX neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
XX Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
XX aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
XX reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
XX vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
XX myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective;  
XX cardiant; cancer; E93 protein.

OS Caenorhabditis elegans.

PN WO200234882-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-US48053.

PR 27-OCT-2000; 2000US-243865P.

PA (UWMA-) UNIV MARYLAND BIOTECHNOLOGY INST.

PI Baehrecke EH;

DR WPI; 2002-479717/51.

XX Novel programmed cell death modulating proteins, useful for treating or  
XX preventing disorders associated with abnormal cell proliferation and  
XX apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
XX infarction -  
XX Claim 1; Fig 1; 88pp; English.

CC The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is nematode E93 programmed cell death modulating protein  
XX conserved domain.

Query Match 78.1%; Score 217; DB 23; Length 53;  
Best Local Similarity 73.6%; Pred. No. 1.6e-23;  
Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

OY 1 KGTREPKRGKRYNYDRDSLVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
DB 1 KRSPRKQYRKRYNDKSLDEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 53

RESULT 5  
AAE24592  
ID AAE24592 standard; Protein; 53 AA.

AC AAE24592;

DT 04-OCT-2002 (first entry)

DE Human E93 programmed cell death modulating protein conserved domain.

XX Human; cancer; programmed cell death modulating protein; adenocarcinoma;  
XX cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
XX neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
XX Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
XX aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
XX reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
XX vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
XX myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective;  
XX cardiant; E93 protein.

OS Homo sapiens.

PN WO200234882-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-US48053.

PR 27-OCT-2000; 2000US-243865P.

PA (UWMA-) UNIV MARYLAND BIOTECHNOLOGY INST.

PI Baehrecke EH;

DR WPI; 2002-479717/51.

XX Novel programmed cell death modulating proteins, useful for treating or  
XX preventing disorders associated with abnormal cell proliferation and  
XX apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
XX infarction -  
XX Claim 1; Fig 1; 88pp; English.

XX The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is human E93 programmed cell death modulating protein conserved  
CC domain.

XX  
SO Sequence 53 AA;  
Query Match 59.4%; Score 165; DB 23; Length 53;  
Best Local Similarity 60.4%; Pred. No. 4.8e-16;  
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNDRLSLVEAVKAVQSGMSVHRASGYGVPHSTLEKYKER 53  
Db 1 KQPKKKRGYRYQDHEIMEEAIAMWMSGKMSVSKAQGIYGVPHSTLEKYKER 53

RESULT 6  
AAE24371 standard; Protein: 442 AA.  
XX  
AC AAE24371;  
XX  
DT 04-OCT-2002 (first entry)  
XX  
DE Human E93 programmed cell death modulating protein.  
XX  
KW Human; cancer; programmed cell death modulating protein; adenocarcinoma;  
KW cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
KW neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
KW Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
KW aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
KW reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
KW vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
KW myeloma; neurotropic; vasotropic; immunostimulant; cerebroprotective;  
KW cardiant; E93 protein.  
XX  
OS Homo sapiens.  
XX  
OS  
XX  
FT Key Location/Qualifiers  
FT Domain 353..405  
FT /note="Conserved domain"  
XX  
XX WO200234882-A2.  
XX  
XX 02-MAY-2002.  
XX  
XX 29-OCT-2001; 2001WO-US48053.  
XX  
XX 27-OCT-2000; 2000US-243865P.  
XX  
XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
XX  
XX Baehrecke EH;  
XX  
XX WPI; 2002-479717/51.  
XX  
XX Novel programmed cell death modulating proteins, useful for treating or

PT preventing disorders associated with abnormal cell proliferation and  
PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
PT infarction -  
XX  
XX Claim 1; Fig 4; 88pp; English.  
XX  
XX The present invention relates to novel programmed cell death modulating  
CC proteins and polynucleotides encoding such proteins. Sequences of the  
CC invention are useful to screen potential cellular apoptosis inhibiting  
CC compounds to determine their use as therapeutic agents for treatment of  
CC diseases associated with increased programmed cell death. They are also  
CC useful for treating or preventing disorders associated with decrease in  
CC apoptosis. Programmed cell death modulating sequences are useful for  
CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
CC invention are useful for treating disorders associated with increase  
CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
CC diseases and other infectious or genetic immunodeficiencies. Sequences  
CC of the invention are used as vaccines and in gene therapy. The present  
CC sequence is human E93 programmed cell death modulating protein.

SO Sequence 442 AA;  
Query Match 59.4%; Score 165; DB 23; Length 442;  
Best Local Similarity 60.4%; Pred. No. 7.2e-15;  
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNDRLSLVEAVKAVQSGMSVHRASGYGVPHSTLEKYKER 53  
Db 353 KQPKKKRGYRYQDHEIMEEAIAMWMSGKMSVSKAQGIYGVPHSTLEKYKER 405

RESULT 7  
ABG17942  
ID ABG17942 standard; Protein: 630 AA.  
XX  
AC ABG17942;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #17933.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Dzmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX N-PSDB; AAS82129.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID NO 48301; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations in  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG0377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 630 AA:

Query Match 59.4%; Score 165; DB 22; Length 630;  
 Best Local Similarity 60.4%; Pred. No. 1.1e-14;  
 Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

OY 1 KGTTPKRGKRYNYDRDLSLEAVKAVQRGEMSVHRAAGSYGVPHSTLEYKVKER 53  
 541 KQPRKKRGYROYDHEIMEAIVAMSGKMSVSKAGIYGVPHSTLEYKVKER 593

RESULT 8  
 ABP32451 ID ABP32451 standard; Protein; 104 AA.

XX AC ABP32451;

XX DT -09-JUL-2002 (first entry)

XX DE Human ORF1424 protein, SEQ ID NO:2848.

XX Human: ORF, open reading frame; ORFX: drug screening; diagnosis;  
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KW immune modulation; haematopoiesis regulation; tissue growth;  
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;  
 KW behaviour; cancer; proliferative disorder; neurological disorder;  
 KW cardiovascular disease; immune system disorder; organ transplantation;  
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KW hypothyroidism; cholesterol ester storage disease; infection; vlnnary;  
 KW vasotrophic; antidiabetic; antidiabetic; cytosolic; neurotropic;  
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 KW caridiatic; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX OS Homo sapiens.

XX PN WO200190366-A2.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-US17076.

XX PR 24-MAY-2000; 2000US-206690P.

XX PA (CURA-) CURAGEN CORP.

XX PI Leach MD, Shimkets RA;

XX WPI; 2002-106200/14.

DR N-PSDB; AEN76477.

XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and disorders related to organ  
 PT transplantation -

XX Claim 10; Page 971-972; 2508pp; English.

CC Sequences ABP31028-ABP35561 represent 4334 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences AEN75054-  
 CC AEN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antifibrotic activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX SQ Sequence 104 AA:

Query Match 58.6%; Score 163; DB 23; Length 104;  
 Best Local Similarity 58.5%; Pred. No. 2.2e-15;  
 Matches 31; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

OY 1 KGTTPKRGKRYNYDRDLSLEAVKAVQRGEMSVHRAAGSYGVPHSTLEYKVKER 53  
 8 KQPRKKRGYROYDHEIMEAIVAMSGKMSVSKAGIYGVPHSTLEYKVKER 60

RESULT 9  
 AAE24594 ID AAE24594 standard; Protein; 53 AA.

XX AC AAE24594;

XX DT 04-OCT-2002 (first entry)

XX DE Mouse E93 programmed cell death modulating protein conserved domain.

XX Mouse; cancer; programmed cell death modulating protein; adenocarcinoma;  
 KW cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
 KW neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
 KW Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
 KW aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
 KW reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
 KW vaccine; gene therapy; lymphoma; cytosolic; melanoma; neuroprotective;  
 KW myeloma; neurotropic; vasotrophic; immunostimulant; ceridroprotective;

KM cardiant; E93 protein.  
 XX Mus musculus.  
 XX WO200234882-A2.  
 XX 02-MAY-2002.  
 XX 29-OCT-2001; 2001WO-US48053.  
 XX 27-OCT-2000; 2000US-243865P.  
 XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
 XX Baehrecke EH;  
 XX WPI; 2002-479717/51.  
 XX  
 XX Novel programmed cell death modulating proteins, useful for treating or  
 PT preventing disorders associated with abnormal cell proliferation and  
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
 PT infarction -  
 XX  
 XX Claim 1; Fig 1; 88bp; English.  
 PS  
 CC The present invention relates to novel programmed cell death modulating  
 CC proteins and polynucleotides encoding such proteins. Sequences of the  
 CC invention are useful to screen potential cellular apoptosis inhibiting  
 CC compounds to determine their use as therapeutic agents for treatment of  
 CC diseases associated with increased programmed cell death. They are also  
 CC useful for treating or preventing disorders associated with decrease in  
 CC apoptosis. Programmed cell death modulating sequences are useful for  
 CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
 CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
 CC invention are useful for treating disorders associated with increase  
 CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
 CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
 CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
 CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
 CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
 CC diseases and other infectious or genetic immunodeficiencies. Sequences  
 CC of the invention are used as vaccines and in gene therapy. The present  
 CC sequence is mouse E93 programmed cell death modulating protein conserved  
 CC domain.  
 CC  
 XX  
 SO Sequence 53 AA;  
 Query Match 56.8%; Score 158; DB 23; Length 53;  
 Best Local Similarity 56.6%; Pred. No. 4.9e-15;  
 Matches 30; Conservative 8; Mismatches 15; Indels 0; Gaps 0;  
 Oy 1 KGTTPKRGKRYNVDRLVEA-VKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 1 KHPKRRGRYKRYNVDRLVEA-VKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 AAE24593  
 ID AAE24593 standard; Protein; 54 AA.  
 XX  
 AC AAE24593;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 XX Fish E93 programmed cell death modulating protein conserved domain.  
 XX  
 XX Fish; cancer; programmed cell death modulating protein; adenocarcinoma;  
 KM cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
 KM neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
 KM Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
 KM aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
 KM reperfusion injury; toxin-induced disease; genetic immunodeficiency;  
 KM vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;

KM myeloma; noctropic; vasotropic; immunostimulant; cerebroprotective;  
 KM cardiant; E93 protein.  
 XX  
 XX Tetraodon nigroviridis.  
 XX WO200234882-A2.  
 XX 02-MAY-2002.  
 XX 29-OCT-2001; 2001WO-US48053.  
 XX 27-OCT-2000; 2000US-243865P.  
 XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
 XX Baehrecke EH;  
 XX WPI; 2002-479717/51.  
 XX  
 XX Novel programmed cell death modulating proteins, useful for treating or  
 PT preventing disorders associated with abnormal cell proliferation and  
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
 PT infarction -  
 XX  
 XX Claim 1; Fig 1; 88bp; English.  
 PS  
 CC The present invention relates to novel programmed cell death modulating  
 CC proteins and polynucleotides encoding such proteins. Sequences of the  
 CC invention are useful to screen potential cellular apoptosis inhibiting  
 CC compounds to determine their use as therapeutic agents for treatment of  
 CC diseases associated with increased programmed cell death. They are also  
 CC useful for treating or preventing disorders associated with decrease in  
 CC apoptosis. Programmed cell death modulating sequences are useful for  
 CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
 CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
 CC invention are useful for treating disorders associated with increase  
 CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
 CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
 CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
 CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
 CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
 CC diseases and other infectious or genetic immunodeficiencies. Sequences  
 CC of the invention are used as vaccines and in gene therapy. The present  
 CC sequence is fish E93 programmed cell death modulating protein conserved  
 CC domain.  
 CC  
 XX  
 SO Sequence 54 AA;  
 Query Match 53.8%; Score 149.5; DB 23; Length 54;  
 Best Local Similarity 59.3%; Pred. No. 8.4e-14;  
 Matches 32; Conservative 4; Mismatches 17; Indels 1; Gaps 1;  
 Oy 1 KGTTPKRGKRYNVDRLVEA-VKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 1 KQPKRRGRYKRYNVDRLVEA-VKAVORGEMSVHRAGSYGVPHSTLEYKVKER 54  
 AAE24373  
 ID AAE24373 standard; Protein; 1221 AA.  
 XX  
 AC AAE24373;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 XX Fruit fly E93 programmed cell death modulating protein #2.  
 XX  
 XX Fruit fly; programmed cell death modulating protein; adenocarcinoma;  
 KM cellular apoptosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
 KM neurodegenerative disease; Alzheimer's disease; retinitis pigmentosa;  
 KM Parkinson's disease; myelodysplastic syndrome; cerebellar degeneration;  
 KM aplastic anaemia; ischaemic injury; myocardial infarction; stroke;  
 KM reperfusion injury; toxin-induced disease; genetic immunodeficiency;

KW vaccine; gene therapy; lymphoma; cytostatic; melanoma; neuroprotective;  
 KW myeloma; nootropic; vasotropic; immunostimulant; cerebroprotective;  
 KW cardiant; cancer; E93 protein.  
 OS Drosophila melanogaster.  
 XX  
 XX WO200234892-A2.  
 PN  
 XX  
 XX 02-MAY-2002.  
 PD  
 XX  
 XX 29-OCT-2001; 2001WO-US48053.  
 PF  
 XX  
 XX 27-OCT-2000; 2000US-243865P.  
 PR  
 XX  
 XX (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.  
 PA  
 XX  
 XX Baehrecke EH;  
 PI  
 XX WPI; 2002-479717/51.  
 DR  
 XX N-PSDB; AAD39238.  
 PT  
 XX Novel programmed cell death modulating proteins, useful for treating or  
 PT preventing disorders associated with abnormal cell proliferation and  
 PT apoptosis such as cancer, stroke, Parkinson's disease, myocardial  
 PT infarction -  
 XX  
 XX PS Disclosure; Page 77-82; 88pp; English.  
 XX  
 XX The present invention relates to novel programmed cell death modulating  
 CC proteins and polynucleotides encoding such proteins. Sequences of the  
 CC invention are useful to screen potential cellular apoptosis inhibiting  
 CC compounds to determine their use as therapeutic agents for treatment of  
 CC diseases associated with increased programmed cell death. They are also  
 CC useful for treating or preventing disorders associated with decrease in  
 CC apoptosis. Programmed cell death modulating sequences are useful for  
 CC treating or preventing cancer e.g. adenocarcinoma, leukaemia, lymphoma,  
 CC melanoma, myeloma. Inhibition of the activity of the sequences of the  
 CC invention are useful for treating disorders associated with increase  
 CC in cell death or apoptosis such as acquired immunodeficiency syndrome  
 CC (AIDS), neurodegenerative diseases (e.g., Alzheimer's disease, retinitis  
 CC pigmentosa, Parkinson's disease and cerebellar degeneration), ischaemic  
 CC injuries (e.g., myocardial infarction, stroke, reperfusion injury),  
 CC myelodysplastic syndromes (e.g., aplastic anaemia), toxin-induced  
 CC diseases and other infectious or genetic immunodeficiencies. Sequences  
 CC of the invention are used as vaccines and in gene therapy. The present  
 CC sequence is fruit fly E93 programmed cell death modulating protein.  
 XX  
 XX SQ Sequence 1221 AA;  
 Query Match 35.6%; Score 99; DB 23; Length 1221;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-05;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KGTTPKRGKRYNYDRDSL 18  
 Db 758 KGTTPKRGKRYNYDRDSL 775  
 RESULT 12  
 ABB67028  
 ID ABB67028 standard; Protein; 1046 AA.  
 XX  
 XX ABB67028;  
 AC  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 27876.  
 DE  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX

PN WO200171042-A2.  
 XX  
 XX PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX  
 XX 23-MAR-2000; 2000US-191637P.  
 PR  
 XX 11-JUL-2000; 2000US-0614150.  
 XX  
 XX PA (PEKE ) PE CORP NY.  
 XX  
 XX PI Venter JC, Adams M, Li PWD, Myers EW,  
 DR WPI; 2001-656860/75.  
 DR  
 XX N-PSDB; ABL11131.  
 PT  
 XX  
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX PS Disclosure; SEQ ID NO 27876; 21pp + Sequence Listing; English.  
 XX  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX SQ Sequence 1046 AA;  
 Query Match 33.3%; Score 92.5; DB 22; Length 1046;  
 Best Local Similarity 35.3%; Pred. No. 0.00057;  
 Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
 OY 3 TRPKRGKRYNYDRDSLVEAVKAVRGEMSVHRAGSYGVPHSTLEKVKER 53  
 Db 753 TRPKRGKRYNYDRDSLVEAVKAVRGEMSVHRAGSYGVPHSTLEKVKER 802  
 RESULT 13  
 ABB59068  
 ID ABB59068 standard; Protein; 1064 AA.  
 XX  
 XX ABB59068;  
 AC  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 3996.  
 DE  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX  
 XX PN WO200171042-A2.  
 XX  
 XX PD 27-SEP-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX  
 XX 23-MAR-2000; 2000US-191637P.  
 PR  
 XX 11-JUL-2000; 2000US-0614150.  
 XX  
 XX PA (PEKE ) PE CORP NY.  
 XX  
 XX PI Venter JC, Adams M, Li PWD, Myers EW,  
 XX

DR WPI; 2001-656860/75.  
DR N-PSDB; ABL03171.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 3996; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX Sequence 1064 AA;  
SQ

Query Match 33.3%; Score 92.5; DB 22; Length 1064;  
Best Local Similarity 35.3%; Pred. No. 0.0059;  
Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
OY 3 TRPKRGKRYNDRDLSLVEAVKAVQKGMVHRAGSYGVPHSTLEYKVKER 53  
Db 771 TPKEGKGTGKSWEDMLQNALHKLHSGQISANKASAKAFIPSTL-YK1ARR 820

RESULT 14  
ABB63113  
ID ABB63113 standard; Protein; 661 AA.  
XX  
XX ABB63113;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 16131.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
XX DR N-PSDB; ABL07216.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 16131; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX Sequence 661 AA;  
SQ

Query Match 30.4%; Score 84.5; DB 22; Length 661;  
Best Local Similarity 37.3%; Pred. No. 0.0045;  
Matches 19; Conservative 13; Mismatches 18; Indels 1; Gaps 1;  
OY 2 GTRPKRGKRYNDRDLSLVEAVKAVQKGMVHRAGSYGVPHSTLEYKVKER 52  
Db 361 GKPEKRYKQYTRADMCALQAVREG-MSALQSRKYGLPRTLYDKVRK 410

RESULT 15  
AAV44303  
ID AAV44303 standard; Protein; 835 AA.  
XX  
XX AAV44303;  
XX  
XX 29-FEB-2000 (first entry)  
XX  
XX Tomato beta galactosidase-1.  
XX  
XX Tomato beta galactosidase-1; TBG; Rutgers tomato plant; pectin;  
XX fruit softening; beta galactosidase II protein; biofilm;  
XX transgenic plant; protoplast isolation.  
XX  
XX Lycopersicon esculentum.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..24  
XX FT label= Signal\_peptide  
XX FT 25..835  
XX FT label= beta-galactosidase-1  
XX  
XX WO9964564-A1.  
XX  
XX 16-DEC-1999.  
XX  
XX 08-JUN-1999; 99WO-US12697.  
XX  
XX 09-JUN-1998; 98US-0088805.  
XX  
XX (USDA) US DEPT OF AGRICULTURE.  
XX  
XX Gross KC, Smith DL;  
XX  
XX WPI; 2000-097532/08.  
XX DR N-PSDB; AA225338.  
XX  
XX New beta-galactosidases, used to prepare transgenic plants with altered  
PT fruit ripening -  
XX  
XX Claim 1; Fig 2; 85pp; English.  
XX  
XX The present sequence is tomato beta galactosidase-1 (TBG-1) encoded by a  
CC cDNA derived from breaker, turning and pink fruit pericarp from 'Rutgers'  
CC tomato plants. This hydrolyses terminal non-reducing beta-D-galactosyl  
CC residues from beta-D-galactosides leading to loss of tissue integrity and  
CC fruit softening. This is used for modifying cell wall metabolism and  
CC controlling ripening of fruit by altering activity of beta galactosidase  
CC II protein. Pectin with reduced galactosyl content is produced for use in  
CC biofilms or solutions. Transgenic plants with altered fruit ripening are  
CC produced by introducing DNA constructs comprising TBG cDNA. TBG forms a  
CC component of an enzyme mixture used to isolate protoplasts.  
XX  
XX Sequence 835 AA;  
SQ

Query Match 23.9%; Score 66.5; DB 21; Length 835;  
Best Local Similarity 42.1%; Pred. No. 2.4;  
Matches 16; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 2 GTRPKRGKYRNYDRDLSLVEAIVKAVQRGEMSVH-RAGSY 38  
| : | | | : | | | : | | | |  
DB 78 GHEPPEGKYFEEERYDLVKFKIKVQEGGLYVHLRIGPY 115

Search completed: October 28, 2003, 12:02:06  
Job time : 10.1374 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 : Search time 4.3899 Seconds  
(without alignments)  
510.826 Million cell updates/sec

Title: US-10-016-768A-1

Perfect score: 278

Sequence: 1 KGTTPKRGKYNRYDRSLVE.....RAGSYGVPHSTLEYKVER 53

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	22.3	202	US-09-026-958-2	Sequence 2, Appli
2	62	22.3	202	US-09-057-860A-2	Sequence 2, Appli
3	62	22.3	202	US-09-390-207-29	Sequence 29, Appli
4	61.5	22.1	349	US-09-252-991A-21699	Sequence 21699, A
5	60.5	21.8	730	US-08-696-944-2	Sequence 2, Appli
6	60.5	21.8	838	US-08-696-944-19	Sequence 19, Appli
7	58	20.9	1385	US-08-687-399-7	Sequence 7, Appli
8	57.5	20.5	100	US-08-160-524A-12	Sequence 12, Appli
9	57	20.5	431	US-08-190-802A-37	Sequence 37, Appli
10	55	19.8	431	US-08-477-346-37	Sequence 37, Appli
11	55	19.8	431	US-08-477-346-37	Sequence 37, Appli
12	55	19.8	431	US-08-477-346-37	Sequence 37, Appli
13	55	19.8	431	US-08-477-346-37	Sequence 37, Appli
14	54	19.4	444	US-09-328-352-6657	Sequence 37, Appli
15	54	19.4	444	US-09-328-352-6657	Sequence 37, Appli
16	54	19.4	724	US-09-562-737-26	Sequence 26, Appli
17	53.5	19.2	284	US-09-134-001C-5083	Sequence 5083, Ap
18	53.5	19.2	294	US-09-328-352-7770	Sequence 7770, Ap
19	53.5	19.2	555	US-08-982-232-14	Sequence 14, Appli
20	53	19.1	219	US-08-771-110-2	Sequence 2, Appli
21	53	19.1	219	US-09-642-000-4	Sequence 4, Appli
22	53	19.1	296	US-09-071-035-40	Sequence 40, Appli
23	53	19.1	317	US-09-071-035-38	Sequence 38, Appli
24	53	19.1	342	US-09-107-532A-4114	Sequence 4114, Ap
25	53	19.1	629	US-09-252-991A-17988	Sequence 17988, A
26	52.5	18.9	433	US-08-417-492-2	Sequence 2, Appli
27	52.5	18.9	727	US-08-482-677-11	Sequence 11, Appli

28	52.5	18.9	727	US-08-650-599A-4	Sequence 4, Appli
29	52.5	18.9	727	US-09-490-517-4	Sequence 4, Appli
30	52.5	18.9	731	US-08-696-944-20	Sequence 20, Appli
31	52	18.7	243	US-09-413-814-6	Sequence 6, Appli
32	52	18.7	1118	US-09-379-523-3	Sequence 3, Appli
33	52	18.7	2584	US-08-936-135-4	Sequence 4, Appli
34	51	18.3	140	US-08-187-780-1	Sequence 1, Appli
35	51	18.3	140	US-08-187-780-1	Sequence 1, Appli
36	51	18.3	140	US-08-478-485-1	Sequence 1, Appli
37	51	18.3	148	US-08-102-691-3	Sequence 3, Appli
38	51	18.3	175	US-08-102-691-2	Sequence 2, Appli
39	51	18.3	177	US-09-019-2	Sequence 8, Appli
40	51	18.3	205	US-09-417-721-8	Sequence 8, Appli
41	51	18.3	205	US-08-417-721-8	Sequence 8, Appli
42	51	18.3	206	US-08-102-691-1	Sequence 1, Appli
43	51	18.3	206	US-08-464-580A-15	Sequence 15, Appli
44	51	18.3	206	US-08-441-629-10	Sequence 10, Appli
45	51	18.3	206	US-08-462-169B-12	Sequence 12, Appli

#### ALIGNMENTS

```
RESULT 1
US-09-026-958-2
; Sequence 2, Application US/09026958
; Patent No. 6150098
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Pacific, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING
; TITLE OF INVENTION: NOVEL SECRETED MAMMALIAN POLYPEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Amgen Inc.
; STREET: One Amgen Center Drive,
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,958
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-522
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-026-958-2

Query Match 22.3%; Score 62; DB 3; Length 202;
Best Local Similarity 38.8%; Pred. No. 0.35;
Matches 19; Conservative 5; Mismatches 13; Indels 12; Gaps 2;

OY 15 RDSLVEAVKAVRGEMGV-----HRAGSYGVPHSTLEYKVE 52
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:
108 RDSLVE-LSPPQGRVSVIFGVASRFVAMSGRGLFVPPFTDCKRKE 155

RESULT 2
US-09-057-860A-2
; Sequence 2, Application US/09057860A
; Patent No. 6277820
```

```

GENERAL INFORMATION:
APPLICANT: Arnon Rosenthal
APPLICANT: Mary Hynes
APPLICANT: Weilan Ye
TITLE OF INVENTION: Method Of Dopaminergic And Serotonergic
TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,860A
FILING DATE: 09-Apr-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1364
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
FAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-057-860A-2

Query Match      22.3%  Score 62;  DB 3;  Length 202;
Best Local Similarity 36.8%;  Pred. No. 0.39;
Matches 19;  Conservative 5;  Mismatches 13;  Indels 12;  Gaps 2;

Qy      15 RDSLVEAVKAVQRGEMSV-----HRAGSYGVPHSTLEYKVE 52
Db      108 RDSLLE-LSPVQRGVVSIFGVASRPFVAMSSRGKLFQVPFFDCKFKE 155

RESULT 3
US-09-390-207-29
Sequence 29, Application US/09390207
Patent No. 6504530
GENERAL INFORMATION:
APPLICANT: Thomason, Arlen
APPLICANT: Liu, Benxian
TITLE OF INVENTION: Fibroblast Growth Factor-Like Polypeptides
FILE REFERENCE: 99-371
CURRENT APPLICATION NUMBER: US/09/390,207
CURRENT FILING DATE: 1999-09-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 29
LENGTH: 202
TYPE: PRT
ORGANISM: Mus musculus
US-09-390-207-29

Query Match      22.3%  Score 62;  DB 4;  Length 202;
Best Local Similarity 38.8%;  Pred. No. 0.39;
Matches 19;  Conservative 5;  Mismatches 13;  Indels 12;  Gaps 2;

Qy      15 RDSLVEAVKAVQRGEMSV-----HRAGSYGVPHSTLEYKVE 52
Db      108 RDSLLE-LSPVQRGVVSIFGVASRPFVAMSSRGKLFQVPFFDCKFKE 155
```

```

RESULT 4
US-09-252-991A-21699
Sequence 21699, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 3142
SEQ ID NO 21699
LENGTH: 349
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21699

Query Match      22.1%  Score 61.5;  DB 4;  Length 349;
Best Local Similarity 31.6%;  Pred. No. 0.93;
Matches 18;  Conservative 12;  Mismatches 22;  Indels 5;  Gaps 1;

Qy      1 KCTPPKRGKRYNDSDLVEAVKAVQRGEMSVHRAGSYG----VPHSTLEYKVE 52
Db      82 EGTQQRGHRDADSVLPVAVGAERGGLAPVLAQGVBEVRVQHHPGDLGTEGVED 138

RESULT 5
US-08-696-944-2
Sequence 2, Application US/08696944
Patent No. 5981831
GENERAL INFORMATION:
APPLICANT: Sumant CHENGAPPA
APPLICANT: Susan A. HELLYER
APPLICANT: John S. REID
TITLE OF INVENTION: No. 5981831e1 Exo-(1-4)-Beta-D Galactanase
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/696,944
FILING DATE: 23-Aug-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/00372
FILING DATE: 23-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9403423.8
FILING DATE: 23-FEB-1994
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-696-944-2

Query Match      21.8%  Score 60.5;  DB 2;  Length 730;
```

Best Local Similarity 42.1%; Pred. No. 3.3;  
Matches 16; Conservative 4; Mismatches 17; Indels 1; Gaps 1;

Oy 2 GTRPRGKYRNYDRSLVEAVKAVQGMYSVH-RAGSY 38  
Db 89 GHEPSFGKYFEDRDLVGFILVQAGLFWHLRIGPF 126

## RESULT 6

US-08-696-944-19  
Sequence 19, Application US/08696944  
Patent No. 5981831

GENERAL INFORMATION:  
APPLICANT: Sumant CHENGAPPA  
APPLICANT: Susan A. HELLYER  
APPLICANT: John S. REID  
APPLICANT: Jacqueline DE SILVA  
TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.

COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MS Word

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/696,944  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/00372  
FILING DATE: 23-FEB-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9403423.8  
FILING DATE: 23-FEB-1994  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 838 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-696-944-19

Query Match 21.8%; Score 60.5; DB 2; Length 838;  
Best Local Similarity 39.5%; Pred. No. 4;  
Matches 15; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

Oy 2 GTRPRGKYRNYDRSLVEAVKAVQGMYSVH-RAGSY 38  
Db 81 GHEPQGGKYFEGRYDLVKFILTIVHAGLVYHLRVGPY 118

US-08-687-399-7  
Sequence 7, Application US/08687399  
Patent No. 5928381

GENERAL INFORMATION:  
APPLICANT: Toft, Annette H.  
APPLICANT: Marcher, Dorte H.  
APPLICANT: Pedersen, Hanne H.  
APPLICANT: Nilsson, Thomas E.  
TITLE OF INVENTION: A Combined Desizing and Bleaching  
TITLE OF INVENTION: Process  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: No. 5928381o No. 5928381disk of No. 5928381th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York

STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,399  
FILING DATE:

CLASSIFICATION: 008  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4127, 204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1385 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-687-399-7

Query Match 20.9%; Score 58; DB 2; Length 1385;  
Best Local Similarity 34.1%; Pred. No. 17;  
Matches 15; Conservative 7; Mismatches 20; Indels 2; Gaps 2;

Oy 3 TRPRGKYRNYDRSLVEAVKAVQGMYSVH-RAGSY 46  
Db 32 TRPARSGRGLNSRILETYRPHL-ELETRASPARG-PHEAL 73

US-08-982-232-7  
Sequence 7, Application US/08982232  
Patent No. 5985570

GENERAL INFORMATION:  
APPLICANT: Amutan, Maria  
APPLICANT: Dunn-Coleman, Nigel  
APPLICANT: Nyssonen, Eini M.  
TITLE OF INVENTION: Identification of and Cloning a Mobile  
TITLE OF INVENTION: Transposon from Aspergillus  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genencor International, Inc.  
STREET: 925 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/982,232  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/703,077  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.

REGISTRATION NUMBER: 33,401  
REFERENCE/DOCKET NUMBER: GC270-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 846-7536

TELEFAX: (415) 845-6504  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 555 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-982-232-7

Query Match 20.7%; Score 57.5; DB 2; Length 555;  
Best Local Similarity 25.9%; Pred. No. 6.4;  
Matches 14; Conservative 16; Mismatches 21; Indels 3; Gaps 2;

QY 1 KGTPEKRGKYNRYDRSLVEAVKAVQGM-SYHRAGSYGVPHSTLEKYKER 53  
DB 4 KASIPSKQVEQEG--ILLAIEAIQKQITSIREARVAVARTTLOARLSGR 55

RESULT 9  
US-08-160-524A-12  
Sequence 12, Application US/08160524A

PATENT No. 5651761  
GENERAL INFORMATION:  
APPLICANT: McAdam, Ruth Anne  
APPLICANT: Dale, Jeremy W.  
APPLICANT: Zainuddin, Zainul Fadziruddin B.  
APPLICANT: Caley, David  
TITLE OF INVENTION: PROBES, KITS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert,  
ADDRESS: Attn: Walter H. Dreger  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,524A  
FILING DATE: 01-DEC-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/752,661  
FILING DATE: 18-OCT-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8903968.9  
FILING DATE: 22-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9000411.0  
FILING DATE: 09-JAN-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB90/00276  
FILING DATE: 22-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-55387-1/WHO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown

US-08-160-524A-12

Query Match 20.5%; Score 57; DB 2; Length 100;  
Best Local Similarity 28.9%; Pred. No. 0.85;  
Matches 11; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

QY 9 KYRNYDRSLVEAVKAVQGM-SYHRAGSYGVPHSTL 46  
DB 4 KTORYSKEFAEAVRTVPEHQLSISGASRLSLPEGTL 41

RESULT 10  
US-08-190-802A-37  
Sequence 37, Application US/08190802A

PATENT No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:  
LENGTH: 431 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CSTF 50kDa, Fig. 20  
US-08-190-802A-37

Query Match 19.8%; Score 55; DB 1; Length 431;  
Best Local Similarity 33.3%; Pred. No. 11;  
Matches 17; Conservative 6; Mismatches 22; Indels 6; Gaps 2;

QY 2 GTRPKRGKYNRYDRSLVEAVKAVQGM-SYHRAGSYGVPHSTL 46  
DB 191 GSRDYTLKLPDYKPSAKRAFKYIQEAMLRSLSFHSGDFIVGTQHPITL 241

RESULT 11  
US-08-477-346-37  
Sequence 37, Application US/08477346

PATENT No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 431 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
-ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CSTF 50kDa, Fig. 20  
US-08-477-346-37

Query Match 19.8%; Score 55; DB 3; Length 431;  
Best Local Similarity 33.3%; Pred. No. 11;  
Matches 17; Conservative 6; Mismatches 22; Indels 6; Gaps 2;

OY 2 GTRPRKRGKRYRVDSDLSVEAVKAVORGEV---SVHRAGSY--YGVPHSTL 46  
Db 191 GSRDYTLKLFDPYKPSAKARAFKYLQEAEMLRISFHPGDFILVGTQHPPTL 241

RESULT 12  
US-08-473-089-37  
Sequence 37, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 431 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
-ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CSTF 50kDa, Fig. 20  
US-08-473-089-37

Query Match 19.8%; Score 55; DB 4; Length 431;  
Best Local Similarity 33.3%; Pred. No. 11;  
Matches 17; Conservative 6; Mismatches 22; Indels 6; Gaps 2;

OY 2 GTRPRKRGKRYRVDSDLSVEAVKAVORGEV---SVHRAGSY--YGVPHSTL 46  
Db 191 GSRDYTLKLFDPYKPSAKARAFKYLQEAEMLRISFHPGDFILVGTQHPPTL 241

RESULT 13  
US-08-487-072A-37  
Sequence 37, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 431 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
-ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CSTF 50kDa, Fig. 20



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:03:24 ; Search time 9.85051 Seconds  
(without alignments)  
901.011 Million cell updates/sec

Title: US-10-016-768A-1

Perfect score: 278  
Sequence: 1 KGTTPKRGKXRNRYDRSLVE.....RAGSYGVPHSTLEYKVER 53

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278	100.0	53	US-10-016-768-1	Sequence 1, Appl1
2	278	100.0	1165	US-10-016-768-10	Sequence 10, Appl1
3	217	78.1	53	US-10-016-768-5	Sequence 5, Appl1
4	165	59.4	53	US-10-016-768-2	Sequence 2, Appl1
5	165	59.4	442	US-10-016-768-8	Sequence 8, Appl1
6	158	56.8	53	US-10-016-768-4	Sequence 4, Appl1
7	149.5	53.8	54	US-10-016-768-3	Sequence 3, Appl1
8	99	35.6	1221	US-10-016-768-11	Sequence 11, Appl1
9	70	25.2	277	US-10-029-386-33895	Sequence 33895, A
10	61	21.9	140	US-09-822-485-33	Sequence 33, Appl1
11	61	21.9	140	US-10-374-207-33	Sequence 33, Appl1
12	61	21.9	162	US-09-822-485-32	Sequence 32, Appl1
13	61	21.9	162	US-10-374-207-32	Sequence 32, Appl1
14	59	21.2	191	US-10-156-761-12095	Sequence 12095, A
15	58.5	21.0	448	US-10-342-224-82	Sequence 82, Appl1

16	58.5	21.0	448	12	US-10-171-404A-20	Sequence 20, Appl1
17	56.5	20.3	673	12	US-09-949-029-100	Sequence 10, Appl1
18	56	20.1	148	9	US-09-822-485-3	Sequence 3, Appl1
19	56	20.1	148	12	US-10-374-207-3	Sequence 3, Appl1
20	56	20.1	170	9	US-09-822-485-2	Sequence 2, Appl1
21	56	20.1	170	9	US-09-750-963-2	Sequence 2, Appl1
22	56	20.1	170	12	US-10-237-496-62	Sequence 62, Appl1
23	56	20.1	170	12	US-10-242-074-62	Sequence 62, Appl1
24	56	20.1	170	12	US-10-242-505-62	Sequence 62, Appl1
25	56	20.1	170	12	US-10-242-574-62	Sequence 62, Appl1
26	56	20.1	170	12	US-10-243-261-62	Sequence 62, Appl1
27	56	20.1	170	12	US-10-243-282-62	Sequence 62, Appl1
28	56	20.1	170	12	US-10-243-402-62	Sequence 62, Appl1
29	56	20.1	170	12	US-10-243-431-62	Sequence 62, Appl1
30	56	20.1	170	12	US-10-245-164-62	Sequence 62, Appl1
31	56	20.1	170	12	US-10-244-572-62	Sequence 62, Appl1
32	56	20.1	170	12	US-10-374-207-2	Sequence 2, Appl1
33	56	20.1	170	12	US-10-197-942-62	Sequence 62, Appl1
34	56	20.1	170	12	US-10-238-196-62	Sequence 62, Appl1
35	56	20.1	170	12	US-10-245-013-62	Sequence 62, Appl1
36	56	20.1	170	12	US-09-998-866-2	Sequence 2, Appl1
37	56	20.1	170	14	US-10-005-646-4	Sequence 4, Appl1
38	56	20.1	170	15	US-10-245-103-62	Sequence 62, Appl1
39	56	20.1	170	15	US-10-245-107-62	Sequence 62, Appl1
40	56	20.1	170	15	US-10-245-143-62	Sequence 62, Appl1
41	56	20.1	170	15	US-10-245-771-62	Sequence 62, Appl1
42	56	20.1	170	15	US-10-245-851-62	Sequence 62, Appl1
43	56	20.1	170	15	US-10-245-883-62	Sequence 62, Appl1
44	56	20.1	170	15	US-10-237-535-62	Sequence 62, Appl1
45	56	20.1	170	15	US-10-238-183-62	Sequence 62, Appl1

## ALIGNMENTS

RESULT 1  
US-10-016-768-1  
; Sequence 1, Application US/10016768  
; Publication No. US2002014243A1  
; GENERAL INFORMATION:  
; APPLICANT: Baehrcke, Eric H.  
; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
; FILE REFERENCE: 4115-131  
; CURRENT APPLICATION NUMBER: US/10/016,768  
; CURRENT FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(54)  
; OTHER INFORMATION: X can be any amino acid  
US-10-016-768-1

Query Match 100.0%; Score 278; DB 14; Length 53;  
Best Local Similarity 100.0%; Pred. No. 9.1e-30;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 KGTTPKRGKXRNRYDRSLVEAVKAVRGEMSVHAGSYGVPHSTLEYKVER 53  
Db 1 KGTTPKRGKXRNRYDRSLVEAVKAVRGEMSVHAGSYGVPHSTLEYKVER 53  
RESULT 2  
US-10-016-768-10  
; Sequence 10, Application US/10016768  
; Publication No. US2002014243A1  
; GENERAL INFORMATION:  
; APPLICANT: Baehrcke, Eric H.  
; TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH

FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 10  
LENGTH: 1165  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-10-016-768-10

Query Match 100.0%; Score 278; DB 14; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 3.8e-28;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
Db 758 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 810

## RESULT 3

US-10-016-768-5  
Sequence 5, Application US/10016768  
Publication No. US20020142443A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 53  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(54)  
OTHER INFORMATION: X CAN BE ANY AMINO ACID  
US-10-016-768-5

Query Match 78.1%; Score 217; DB 14; Length 53;  
Best Local Similarity 73.6%; Pred. No. 1.2e-21;  
Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
Db 1 KRSPKRGQRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53

## RESULT 4

US-10-016-768-2  
Sequence 2, Application US/10016768  
Publication No. US20020142443A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 53  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(54)  
OTHER INFORMATION: X CAN BE ANY AMINO ACID  
US-10-016-768-2

Query Match 59.4%; Score 165; DB 14; Length 53;  
Best Local Similarity 60.4%; Pred. No. 1e-14;  
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
Db 1 KQPRKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53

## RESULT 5

US-10-016-768-8  
Sequence 8, Application US/10016768  
Publication No. US20020142443A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 442  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-016-768-8

Query Match 59.4%; Score 165; DB 14; Length 442;  
Best Local Similarity 60.4%; Pred. No. 1.3e-13;  
Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
Db 353 KQPRKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 405

## RESULT 6

US-10-016-768-4  
Sequence 4, Application US/10016768  
Publication No. US20020142443A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 53  
TYPE: PRT  
ORGANISM: M. musculus  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(54)  
OTHER INFORMATION: X can be any amino acid  
US-10-016-768-4

Query Match 56.8%; Score 158; DB 14; Length 53;  
Best Local Similarity 56.6%; Pred. No. 8.8e-14;  
Matches 30; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Oy 1 KGTTPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53  
Db 1 KHPKRGKRYNRYDRDSLVEAVKAVQSGMSVHRAGSYGVPHSTLEYKVKER 53

## RESULT 7

US-10-016-768-3  
Sequence 3, Application US/10016768  
Publication No. US20020142443A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.

TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 3  
LENGTH: 54  
TYPE: PRT  
ORGANISM: T. nigroviridis  
US-10-016-768-3

Query Match 53.8%; Score 149.5; DB 14; Length 54;  
Best Local Similarity 59.3%; Pred. No. 1.2e-12;  
Matches 32; Conservative 4; Mismatches 17; Indels 1; Gaps 1;

QY 1 KGTTPKRGKRYNYPDRSLVE-AVKAVORGEMSVHAGSYGVPHSTLEYKVKER 53  
DB 1 KQPRKRGKRYROYDHDLEASITVMAGRMSVSAQGVGTGIPHSTLEYKVKER 54

RESULT 8  
US-10-016-768-11  
Sequence 11, Application US/10016768  
Publication No. US2002014243A1  
GENERAL INFORMATION:  
APPLICANT: Baehrcke, Eric H.  
TITLE OF INVENTION: GENES REGULATING PROGRAMMED CELL DEATH  
FILE REFERENCE: 4115-131  
CURRENT APPLICATION NUMBER: US/10/016,768  
CURRENT FILING DATE: 2001-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 11  
LENGTH: 1221  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-10-016-768-11

Query Match 35.6%; Score 99; DB 14; Length 1221;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGTTPKRGKRYNYPDRSL 18  
DB 758 KGTTPKRGKRYNYPDRSL 775

RESULT 9  
US-10-029-386-33895  
Sequence 33895, Application US/10029386  
Publication No. US2003019470A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
FILE REFERENCE: AEWICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33895  
LENGTH: 277  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ACO05768.16  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.85  
OTHER INFORMATION: SWISSPROT HIT: Q9Y1D8, EVALU 1.60e+00  
US-10-029-386-33895

Query Match 25.2%; Score 70; DB 12; Length 277;  
Best Local Similarity 41.4%; Pred. No. 0.34;  
Matches 12; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 18 LVEAVKAVORGEMSVHAGSYGVPHSTL 46  
DB 21 LSKALKDIOGALDINKAGILYGIPOKTL 49

RESULT 10  
US-09-822-485-33  
Sequence 33, Application US/09822485  
Patent No. US2002001825A1  
GENERAL INFORMATION:  
APPLICANT: Itoh, No. US2002001825A1  
TITLE OF INVENTION: No. US2002001825A1  
FILE REFERENCE: 08035.0001-01000  
CURRENT APPLICATION NUMBER: US/09/822,485  
CURRENT FILING DATE: 2001-04-02  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 33  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus sp.  
US-09-822-485-33

Query Match 21.9%; Score 61; DB 9; Length 140;  
Best Local Similarity 30.2%; Pred. No. 2.4;  
Matches 19; Conservative 13; Mismatches 15; Indels 16; Gaps 3;

QY 1 KGTTPKRGKRYNYPDRSLVE-AVKAVORG-EMSVHAGSYGVPHSTLEYKV 50  
DB 31 QGTTPKRGKRYNYPDRSLVE-AVKAVORG-EMSVHAGSYGVPHSTLEYKV 50

QY 51 KER 53  
DB 85 RER 87

RESULT 11  
US-10-374-207-33  
Sequence 33, Application US/10374207  
Publication No. US20030170822A1  
GENERAL INFORMATION:  
APPLICANT: Itoh, No. US20030170822A1  
TITLE OF INVENTION: Fibroblast Growth Factor-Like Molecules and Uses Thereof  
FILE REFERENCE: 08035.0001-02000  
CURRENT APPLICATION NUMBER: US/10/374,207  
CURRENT FILING DATE: 2003-07-25  
PRIOR APPLICATION NUMBER: US 09/822,485  
PRIOR FILING DATE: 2001-04-02  
PRIOR APPLICATION NUMBER: US 09/540,118  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 33  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Mus sp.  
US-10-374-207-33

Query Match 21.9%; Score 61; DB 12; Length 140;  
Best Local Similarity 30.2%; Pred. No. 2.4;  
Matches 19; Conservative 13; Mismatches 15; Indels 16; Gaps 3;

QY 1 KGTTPKRGKRYNYPDRSLVE-AVKAVORG-EMSVHAGSYGVPHSTLEYKV 50  
DB 31 QGTTPKRGKRYNYPDRSLVE-AVKAVORG-EMSVHAGSYGVPHSTLEYKV 50  
QY 51 KER 53  
DB 85 RER 87

RESULT 12  
US-09-822-485-32  
; Sequence 32, Application US/09822485  
; Patent No. US2002001825A1  
; GENERAL INFORMATION:  
; APPLICANT: Itoh, No. US20020001825A1uyuki Fibroblast Growth Factor-Like Polypeptides  
; TITLE OF INVENTION: No. US20020001825A1el  
; FILE REFERENCE: 08035.0001-01000  
; CURRENT APPLICATION NUMBER: US/09/822,485  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 32  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-822-485-32

Query Match 21.9%; Score 61; DB 9; Length 162;  
Best Local Similarity 30.2%; Pred. No. 2.8; Indels 16; Gaps 3;  
Matches 19; Conservative 13; Mismatches 15;

OY 1 KGTBPKRGKRYNDRDLSVE-----AVKAVORG-EMSVHRAGSYGVPHSTLEKXV 50  
Db 53 QGTWRHG-----QDSIVEIRSVRGTVVIKAVYSGFYVAMHRRGRGXGSRVYSVDCRF 106  
OY 51 KER 53  
Db 107 RER 109

RESULT 13  
US-10-374-207-32  
; Sequence 32, Application US/10374207  
; Publication No. US20030170822A1  
; GENERAL INFORMATION:  
; APPLICANT: Itoh, No. US20030170822A1uyuki  
; TITLE OF INVENTION: Fibroblast Growth Factor-Like Molecules and Uses Thereof  
; FILE REFERENCE: 08035.0001-02000  
; CURRENT APPLICATION NUMBER: US/10/374,207  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR APPLICATION NUMBER: US 09/822,485  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: US 09/540,118  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 32  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-374-207-32

Query Match 21.9%; Score 61; DB 12; Length 162;  
Best Local Similarity 30.2%; Pred. No. 2.8; Indels 16; Gaps 3;  
Matches 19; Conservative 13; Mismatches 15;

OY 1 KGTBPKRGKRYNDRDLSVE-----AVKAVORG-EMSVHRAGSYGVPHSTLEKXV 50  
Db 53 QGTWRHG-----QDSIVEIRSVRGTVVIKAVYSGFYVAMHRRGRGXGSRVYSVDCRF 106  
OY 51 KER 53  
Db 107 RER 109

RESULT 14  
US-10-156-761-12095  
; Sequence 12095, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 12095  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-12095

Query Match 21.2%; Score 59; DB 15; Length 191;  
Best Local Similarity 40.7%; Pred. No. 6.4;  
Matches 11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 6 KRGKRYNDRDLSVEAVKAVORGEMSV 32  
Db 121 EGGAYDQLERDSLTKAMKGLRRQREV 147

RESULT 15  
US-10-342-224-82  
; Sequence 82, Application US/10342224  
; Publication No. US20030162294A1  
; GENERAL INFORMATION:  
; APPLICANT: Nathalie Verbuggen  
; TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress  
; FILE REFERENCE: CN-01205  
; CURRENT APPLICATION NUMBER: US/10/342,224  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US/09/762,154  
; PRIOR FILING DATE: 2002-02-02  
; PRIOR APPLICATION NUMBER: EP 98202634.6  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 82  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-342-224-82

Query Match 21.0%; Score 58.5; DB 12; Length 448;  
Best Local Similarity 32.0%; Pred. No. 21;  
Matches 16; Conservative 9; Mismatches 18; Indels 7; Gaps 2;

OY 3 TRPRGKRYNDRDLSVE-----AVKAVORGEMSVHRAGSY--YGVPHST 45  
Db 201 TAEKVGEXKYTYDKAVEARDYTAEKAIKAKDKTAERTGEXKYTYVEKAT 250

Search completed: October 28, 2003, 12:17:00  
Job time : 11.8505 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 3.21212 Seconds  
(without alignments)  
1586.783 Million cell updates/sec

Title: US-10-016-768A-1

Perfect score: 278

Sequence: 1 KGTREPKRGKYNRYDRSLVE.....RAGSYGVPHSTLEKVKER 53

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	78.1	185	2 T24276	hypothetical prote
2	99	35.6	1221	2 T13283	probable transcrip
3	92.5	33.3	1085	2 S66149	gene pipsqueak pro
4	67.5	24.3	158	2 A69178	conserved hypothet
5	66.5	23.9	835	2 T06590	probable beta-gala
6	64.5	23.2	729	2 T04269	probable beta-gala
7	64	23.0	188	2 D64176	hypothetical prote
8	64	23.0	753	2 A27041	lysine kinase-re
9	63	22.7	368	2 C90487	oxidoreductase [im
10	62	22.3	202	1 TVMSHS	fibroblast growth
11	61.5	22.1	378	2 E96724	hypothetical prote
12	60	21.6	387	2 G90359	conserved hypothet
13	59	21.2	190	2 T35381	probable RNA polym
14	59	21.2	663	2 B87499	glycosyl transfera
15	58.5	21.0	448	2 H84782	late embryogenesis
16	58.5	21.0	448	2 JC6171	embryonic protein
17	58.5	21.0	555	2 S04909	probable 4-ALPHA-G
18	58.5	21.0	724	2 G70928	beta-galactosidase
19	58.5	21.0	724	2 T04340	probable carbonic
20	58	20.9	286	2 B96615	hypothetical prote
21	57.5	20.7	160	2 G70467	murine hydrolase I
22	57.5	20.7	406	2 AG3021	hypothetical prote
23	57.5	20.7	406	2 C98263	hypothetical prote
24	57	20.5	100	2 S03411	hypothetical prote
25	57	20.5	108	2 S77752	probable phosphor
26	57	20.5	601	2 T21329	hypothetical prote
27	56.5	20.3	1880	2 T18531	tractin - medica
28	56.5	20.3	434	2 H70013	conserved hypothet
29	56.5	20.3	513	2 T10830	nitrogenase [EC 1.

30	56.5	20.3	535	2 S51577	transposase - rice
31	56.5	20.3	2606	2 T24157	hypothetical prote
32	56	20.1	100	2 T44485	conserved hypothet
33	56	20.1	113	2 D85655	unknown in IS [imp
34	56	20.1	113	2 B90794	hypothetical prote
35	56	20.1	421	2 AG2587	lytic murein trans
36	56	20.1	421	2 G97369	hypothetical prote
37	55.5	20.0	434	2 JU0182	monodehydroascorba
38	55.5	20.0	737	2 S44862	ROSD3.2 protein -
39	55	19.8	236	2 T19835	transcription acti
40	55	19.8	236	2 D84103	two-component resp
41	55	19.8	430	2 F97266	aspartyl-tRNA synt
42	55	19.8	431	2 A45142	cleavage stimulat
43	55	19.8	463	2 S27757	embryonic abundant
44	55	19.8	532	2 G70536	probable csp prot
45	54.5	19.6	207	2 C72223	guanylate kinase -

#### ALIGNMENTS

RESULT 1  
T24276  
hypothetical protein T01C1.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T24276  
R:Lenhard, N.  
Submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19868  
A:Accession: T24276  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-185 <M1>  
A:Cross-references: EMBL:Z68010; PIDN:CAA92009.1; GSPDB:GN00028; CESP:T01C1.3  
A:Experimental source: clone T01C1  
C:Genetics:  
A:Gene: CESP:T01C1.3  
A:Map position: X  
A:Introns: 25/3; 93/2; 131/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein T01C1.3

Query Match  
Best Local Similarity 78.1%; Score 217; DB 2; Length 185;  
Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

QY 1 KGTREPKRGKYNRYDRSLVEAVKAVQSGMSVHRAGSYGVPHSTLEKVKER 53  
DB 83 KSRPRKRGQYKRYDKNALDEAVRSVRGEMTVHRAGSFVGVPHTLEKVKER 135

RESULT 2  
T13283  
probable transcription factor E93 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C:Accession: T13283  
R:Baehrecke, E.H.; Thummel, C.S.  
Dev. Biol. 171, 85-97, 1995  
A>Title: The Drosophila E93 gene from the 93F early puff displays stage- and tissue-spec  
A:Reference number: Z17648; MUID:96018744; PMID:7556910  
A:Accession: T13283  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1221 <BAE>  
A:Cross-references: EMBL:U25686; NID:9886047; PID:9886048; PIDN:AAA83228.1  
A:Experimental source: strain Canton S  
C:Genetics:  
A:Gene: E93  
A:Cross-references: FlyBase:FBgn0013948  
A:Map position: 3R  
C:Function:  
A:Description: probably acts in a stage-specific regulatory hierarchy in the salivary gl.

Query Match 35.6%; Score 99; DB 2; Length 1221;  
A:Cross-references: GB:AE000841; GB:AE000666; NID:92621665; PIDN:AB85095.1; PID:92621666  
Best Local Similarity 100.0%; Pred. No. 0.00064;  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH589  
A:Start codon: GTG  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KGTREKRGKYNVDRDSL 18  
Db 758 KGTREKRGKYNVDRDSL 775

RESULT 3  
666149  
gene pipsqueak protein A long form - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 23-Sep-2002  
C:Accession: S66149; S66150; T45461  
R:Weber, U.; Siegel, V.; Mlodzik, M.  
EMBO J. 14, 6247-6257, 1995  
A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for D  
A:Reference number: S66148; MUID:96134923; PMID:8557044  
A:Accession: S66149  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1085 <WEB>  
A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62474.1; PID:g1149500  
A:Accession: S66150  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 'MQ', 428-1085 <WE2>  
A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62475.1; PID:g1149501  
R:Horowitz, H.; Berg, C.A.  
Development 122, 1859-1871, 1996  
A:Title: The Drosophila pipsqueak gene encodes a nuclear BTR-domain-containing protein X  
A:Reference number: Z22972; MUID:96232300; PMID:8674425  
A:Accession: T45461  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-355, 'E', 357-1005, 'H', 1007-1020, 'Q', 1021-1061, 'ERS' <HOR>  
A:Cross-references: EMBL:U48358; NID:g1203906; PIDN:AA647153.1; PID:g1203907  
A:Experimental source: tissue type ovarian  
C:Genetics:  
A:Gene: pipsqueak; psq  
A:Map position: 11  
A:Introns: 427/3  
C:Function:  
A:Description: required for establishing polarity of the developing egg chamber  
C:Superfamily: Broeze-2 protein; POZ domain homology  
F:21-123/Domain: POZ domain homology <POZ>

Query Match 33.3%; Score 92.5; DB 2; Length 1085;  
Best Local Similarity 35.3%; Pred. No. 0.0036;  
Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;

Qy 3 TRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSYGVPHSTLEYKVKR 53  
Db 771 TRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSYGVPHSTLEYKVKR 820

RESULT 4  
A69178  
conserved hypothetical protein MTH589 - Methanobacterium thermoautotrophicum (strain Del  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: A69178  
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vitacek, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: A69178  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-158 <MTH>  
A:Cross-references: GB:AE000841; GB:AE000666; NID:92621665; PIDN:AB85095.1; PID:92621666  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH589  
A:Start codon: GTG  
Matches 16; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

Query Match 24.3%; Score 67.5; DB 2; Length 158;  
Best Local Similarity 34.8%; Pred. No. 0.57;  
Matches 16; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

Qy 6 KRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSYGVPHSTLEYKVKR 51  
Db 112 KRGKYNVDRSLTRR-VREARNMGPAKSKDGLRLRYVYLK 156

RESULT 5  
T06590  
probable beta-galactosidase (EC 3.2.1.23) - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 19-May-2000  
C:Accession: T06590  
R:Carey, A.T.; Holt, K.; Picard, S.; Wilde, R.; Tucker, G.A.; Bird, C.R.; Schuch, W.; Sey  
Plant Physiol. 108, 1099-1107, 1995  
A:Title: Tomato exo-(1-4)-beta-D-galactanase: isolation, changes during ripening in nor  
A:Reference number: Z15780; MUID:95357407; PMID:7630937  
A:Accession: T06590  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-835 <CAR>  
A:Cross-references: EMBL:X83854; NID:9971484; PIDN:CAA58734.1; PID:9971485  
A:Experimental source: cultivar Alisa Craig; pericarp  
C:Superfamily: beta-galactosidase bga  
C:Keywords: glycosidase; hydrolase

Query Match 23.9%; Score 66.5; DB 2; Length 835;  
Best Local Similarity 42.1%; Pred. No. 4.6;  
Matches 16; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

Qy 2 GTRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSY 38  
Db 78 GTRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSY 115

RESULT 6  
T04269  
probable beta-galactosidase (EC 3.2.1.23) - Arabidopsis thaliana  
N:Alternate names: protein P20B18.250  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 19-May-2000  
C:Accession: T04269  
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hehlisel, J.; Mewes, H.W.; Mayer, K.F.X  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z15263  
A:Accession: T04269  
A:Molecule type: DNA  
A:Residues: 1-729 <BEV>  
A:Cross-references: EMBL:AL049483  
A:Experimental source: cultivar Columbia; BAC clone P20B18  
C:Genetics:  
A:Map position: 4  
A:Introns: 58/3; 90/3; 128/2; 150/3; 181/3; 229/3; 259/2; 294/3; 323/1; 362/3; 416/3; 477/  
A>Note: P20B18.250  
C:Superfamily: beta-galactosidase bga  
C:Keywords: glycosidase; hydrolase

Query Match 23.2%; Score 64.5; DB 2; Length 729;  
Best Local Similarity 42.1%; Pred. No. 7;  
Matches 16; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

Qy 2 GTRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSY 38  
Db 83 GTRPKRGKYNVDRDSLVEAVKAVQRCGMSVVRAGSY 120

## RESULT 7

D64176  
 hypothetical protein H11720 - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 21-Jul-2000  
 C:Accession: D64176  
 R:Fejlschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Smal, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630; PMID:7542800  
 A:Accession: D64176  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-188 <TIGR>  
 A:Cross-references: GB:U3845; GB:I42023; NID:G3212236; PIDN:AA23366.1; PID:G1574576; T A:Note: best homolog was a hypothetical protein (insertion element IS1223) from Lactobac

## Query Match

23.0%; Score 64; DB 2; Length 188;  
 Best Local Similarity 32.4%; Pred. No. 1.9;  
 Matches 12; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

OY 8 GKYRNYDRSLVEAVKAVORGEMSVHRAGSYGVPHSTL 44

DB 75 GKKNYSPEFKLVNIOAVKNGKFSAEACLFHFIANS 111

## RESULT 8

A27041  
 tyrosine kinase-related protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 04-Feb-2000  
 C:Accession: A27041  
 R:Haller, J.; Cole, S.; Broemer, G.; Jaekle, H.  
 Genes Dev. 1, 862-867, 1987

A:Title: Dorsal and neural expression of a tyrosine kinase-related Drosophila gene during  
 A:Reference number: A27041; MUID:88112827; PMID:3428600  
 A:Accession: A27041  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-753 <HAL>

C:Genetics:

A:Gene: dtkr  
 A:Cross-references: FlyBase:FBgn0003715

A:Map:position: 2R, 60F1

A:introns: 453/1; 497/1  
 C:Keywords: autophosphorylation; glycoprotein; phosphoprotein  
 F:9,65,187,223,224,250,61,660/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:744/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

## Query Match

Best Local Similarity 23.0%; Score 64; DB 2; Length 753;  
 Matches 12; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

OY 5 PRGKRYNDRSLVEAVKAVORGEMSVHRAGSYGVPHSTL 46

DB 507 PRGKPRSWTNTLTETALQHVNNKMTTSQASRIFGIPYNL 548

## RESULT 9

C90487  
 oxidoreductase [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001

C:Accession: C90487

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-aret, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: C90487

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <KUN>

C:Genetics:

A:Cross-references: GB:AE006641; NID:G13816456; PIDN:AAK43154.1; GSPDB:GN00155

A:Gene: SSO3054

C:Superfamily: fission yeast pyridoxine 4-dehydrogenase

## Query Match

Best Local Similarity 22.7%; Score 63; DB 2; Length 368;  
 Matches 16; Conservative 9; Mismatches 13; Indels 4; Gaps 2;

OY 11 RNYDRSLVEAVKAVORGEMSVHRAGSYGVPHSTLKYKE 52

DB 104 KQYDRSLVLAATKV--RGKMAEHANGE--GISRKIMQVRE 141

## RESULT 10

TWMSHS  
 fibroblast growth factor 4 - mouse

M:Alternate names: transforming protein hctf1, transforming protein k-FGF; transforming  
 C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 17-Mar-2000

C:Accession: S04741; A37360

R:Brooker, S.; Smith, R.; Thurlow, J.; Dickson, C.; Peters, G.

Nucleic Acids Res. 17, 4037-4045, 1989

A:Title: The mouse homologue of hsc/k-FGF: sequence, genome organization and location re

A:Reference number: S04741; MUID:89296455; PMID:2740210

A:Accession: S04741

A:Molecule type: DNA

A:Residues: 1-202 <BRO>

A:Cross-references: GB:X14849; GB:M28516; NID:G52791; PIDN:CAA32967.1; PID:G52792

R:Hebert, J.M.; Basilico, C.; Goldfarb, M.; Haub, O.; Martin, G.R.

Dev. Biol. 138, 454-463, 1990

A:Title: Isolation of cDNAs encoding four mouse FGF family members and characterization

A:Reference number: A37360; MUID:90201563; PMID:2318343

A:Accession: A37360

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-166, 'S', 168-202 <HEB>

A:Cross-references: GB:M30642; NID:G193290; PIDN:AAA37619.1; PID:G309237

C:Genetics:

A:Gene: hsc

C:Superfamily: fibroblast growth factor

C:Keywords: growth factor; transforming protein

## Query Match

Best Local Similarity 22.3%; Score 62; DB 1; Length 202;  
 Matches 19; Conservative 5; Mismatches 13; Indels 12; Gaps 2;

## RESULT 11

E96724  
 hypothetical protein F20P5.11 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: E96724

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federespiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malci, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: E86724  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-378 <STO>  
A:Cross-references: GB:AE005173; NID:g2194124; PIDN:AB61099.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F20P5.11  
A:Map position: 1

Query Match 22.1%; Score 61.5; DB 2; Length 378;  
Best Local Similarity 27.3%; Pred. No. 8.1;  
Matches 15; Conservative 9; Mismatches 22; Indels 9; Gaps 1;

Oy 5 PKRGKYYR-----NYDRSLAEAVKAVRGEMSVHRAGSYGVPHSTLEKYK 50  
Db 243 PPSGKFHDADEBNVWVSGDLDSFSLVTAADLESVAVHEIGHLGLGHSVEESI 297

RESULT 12  
G90359  
conserved hypothetical protein [imported] - *Sulfolobus solfataricus*  
C:Species: *Sulfolobus solfataricus*  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001  
C:Accession: G90359  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.D.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: *Sulfolobus solfataricus* complete genome.  
A:Reference number: A89139  
A:Accession: G90359  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-387 <KUR>  
A:Cross-references: GB:AE006641; NID:g1315214; PIDN:AAK42134.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SSO1939  
C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1618

Query Match 21.6%; Score 60; DB 2; Length 387;  
Best Local Similarity 32.7%; Pred. No. 13;  
Matches 18; Conservative 9; Mismatches 18; Indels 10; Gaps 3;

Oy 3 TRPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEKYK 52  
Db 5 TRPKEDRKLDYDEKEIMIKDSIRAGEMIAVLGMRRIGK-----TSVVNVAVKE 54

RESULT 13  
T35381  
probable RNA polymerase sigma factor - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 04-Mar-2000  
C:Accession: T35381  
R:Murphy, L.; Harris, D.; James, K.D.; Pakkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1999  
A:Reference number: 221576  
A:Accession: T35381  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-190 <MUR>  
A:Cross-references: EMBL:AL079348; PIDN:CAB45480.1; GSPDB:GN00070; SCOEDB:SC66T3.24C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC66T3.24C  
C:Superfamily: transcription initiation factor sigma E

Query Match 21.2%; Score 59; DB 2; Length 190;  
Best Local Similarity 40.7%; Pred. No. 7.9;  
Matches 11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 6 KRGRYRNDRDLSLVEAVKAVRGEMSV 32  
Db 120 EBGAYDQLERDLSLKANKGLOROREV 146

RESULT 14  
E87499  
glycosyl transferase family protein CC2018 [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: E87499  
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eissen, J.; Heidelberg, J.L.  
B.; Lamb, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolome  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: E87499  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-663 <STO>  
A:Cross-references: GB:AE005673; NID:g13423491; PIDN:AAK3993.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2018

Query Match 21.2%; Score 59; DB 2; Length 663;  
Best Local Similarity 32.1%; Pred. No. 30;  
Matches 18; Conservative 13; Mismatches 17; Indels 8; Gaps 3;

Oy 2 GTRPKRGKRYNDRDLSLVEAVKAVRGEMSVHRA---GSYGVPHSTLEKYK 53  
Db 336 GPKRFGGEVWSH--DALESAL--LRGGSVHLAPYDGSYEBSRSLDFPARRDR 387

RESULT 15  
H84782  
late embryogenesis abundant protein (ATECP63) [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
C:Accession: H84782  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84782  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-448 <STO>  
A:Cross-references: GB:AE002093; NID:g4415909; PIDN:AAD20140.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g36640  
A:Map position: 2  
C:Superfamily: pea seed biotin-containing protein

Query Match 21.0%; Score 58.5; DB 2; Length 448;  
Best Local Similarity 32.0%; Pred. No. 23;  
Matches 16; Conservative 9; Mismatches 18; Indels 7; Gaps 2;

Oy 3 TRPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSY--YGVPHST 45  
Db 201 TAEKVGKDYDVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAVKAT 250

Search completed: October 28, 2003, 12:03:11  
Job time : 8.21212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 1.92727 Seconds  
(without alignments)  
1293.234 Million cell updates/sec

Title: US-10-016-768a-1

Perfect score: 278  
Sequence: 1 KGRPKRGKYNRYDRSLVE.....RAGSYGVPHSTLEYKXER 53

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.5	23.9	835	1	BGAL_LYCSES
2	64	23.0	188	1	YH20_HAEIN
3	64	22.0	753	1	TKR_DROME
4	62	22.3	202	1	FCF4_MOUSE
5	58.5	21.0	555	1	LED8_DAUCA
6	58.5	21.0	724	1	MALQ_MYCTU
7	57.5	20.7	160	1	Y059_AQUAE
8	57	20.5	100	1	Y1S1_SHISO
9	56.5	20.3	128	1	YHUA_SCHPO
10	56.5	20.3	513	1	NIRK_RHISN
11	56	20.1	170	1	FCFM_HUMAN
12	56	20.1	977	1	BAB1_DROME
13	55.5	20.0	737	1	YNC2_CAEEL
14	55	19.8	236	1	DEGU_BACBR
15	55	19.8	431	1	CST1_HUMAN
16	54.5	19.6	207	1	KGUA_THEMA
17	54.5	19.6	384	1	YFJ3_GULSO
18	54	19.4	162	1	FCFM_MOUSE
19	54	19.4	1716	1	RPA1_RAT
20	54	19.4	1717	1	RPA1_MOUSE
21	53.5	19.2	494	1	VPE_CITSI
22	53.5	19.2	511	1	YEOB_YEAST
23	53.5	19.2	914	1	IFP2_CITTE
24	53	19.1	398	1	ARGD_METUA
25	53	19.1	868	1	MCW2_YEAST
26	52.5	18.9	727	1	NETA_DROME
27	52.5	18.9	731	1	BGAL_DIACA
28	52	18.7	194	1	ORN_XANCP
29	52	18.7	309	1	IFRH_MAIZE
30	52	18.7	454	1	CTR1_HUMAN
31	52	18.7	730	1	SEC6_SCHPO
32	52	18.7	1067	1	BAB2_DROME
33	51.5	18.5	211	1	KGUA_STRP3

34	51.5	18.5	211	1	KGUA_STRPY	099ym5 streptococc
35	51.5	18.5	267	1	IF2A_ARCFU	029723 archaeoglob
36	51.5	18.5	398	1	TRMU_AGRFS	08um95 agrobacteri
37	51.5	18.5	398	1	TRMU_BRUME	08y116 bruceella me
38	51.5	18.5	398	1	TRMU_RHIME	092mb5 rhizobium m
39	51.5	18.5	473	1	SYC_METAC	08tsp6 methanosarc
40	51.5	18.5	538	1	PME2_CAEEL	009525 caenorhabdi
41	51.5	18.5	579	1	URA7_YEAST	P28274 saccharomyc
42	51	18.3	206	1	FCF4_HUMAN	P08620 homo sapien
43	51	18.3	208	1	HCFE_SEROF	P56825 sepiia offic
44	51	18.3	216	1	R10A_ICTPU	090yv8 ictalurus p
45	51	18.3	224	1	MTGA_ACTICA	024849 actinobact

## ALIGNMENTS

RESULT 1  
ID BGAL\_LYCSES STANDARD: PRT; 835 AA.  
AC P48980:  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-galactosidase) (Exo-(1->4)-beta-D-galactanase).  
OS Lycopodium obscurum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=cv. Alisa Craig; TISSUE=pericarp;  
RX MEDLINE=95357407; PubMed=7630937;  
RA Carey A.T., Holt K., Picard S., Wilde R., Tucker G.A., Bird C.R., RA Schuch W., Seymour G.B.;  
RT "Tomato exo-(1->4)-beta-D-galactanase. Isolation, changes during ripening in normal and mutant tomato fruit, and characterization of a related cDNA clone.";  
RT Plant Physiol. 108:1099-1107(1995).  
RL -!- FUNCTION: Involved in cell wall degradation. Degrades polysaccharides containing beta-(1->4)-linked galactans, acting as an exo-(1->4)-beta-D-galactanase.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-galactose residues in beta-D-galactosides.  
CC -!- MISCELLANEOUS: Has a pH optimum of 4.5.  
CC -!- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.  
CC -!- SIMILARITY: Contains 1 SUEL-type lectin domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: X83654; CAA58734.1; -  
CC PIR: T06590; T06590.  
CC InterPro: IPR000922; Gal lectin.  
DR InterPro: IPR001944; Glyco\_hydro\_35.  
DR Pfam: PF02140; Gal\_Lectin; 1.  
DR Pfam: PF01301; Glyco\_hydro\_35; 1.  
DR PRINTS: PR00742; GLHYDRLASE35.  
DR ProDom: PD005612; Gal lectin; 1.  
DR PROSITE: PS01182; GLYCOSYL\_HYDROL\_F35; 1.  
DR PROSITE: PS50228; SUEL\_LECTIN; 1.  
KW Hydrolyase; Glycosidase; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 835 BETA-GALACTOSIDASE.  
FT DOMAIN 749 835 SUEL-TYPE LECTIN.  
FT FT  
FT ACT\_SITE 180 180 PROTON DONOR (POTENTIAL).

FT ACT SITE 249 249 NUCLEOPHILE (POTENTIAL).  
 SQ SEQUENCE 835 AA; 93336 MW; 94C9685F95CA4646 CRC64;  
 Query Match 23.9%; Score 66.5; DB 1; Length 835;  
 Best Local Similarity 42.1%; Pred. No. 2.1;  
 Matches 16; Conservative 5; Mismatches 16; Indels 1; Gaps 1;  
 Oy 2 GTRPRGRKRYNDRLSLVEAVKAVRGEMSVH-RAGSY 38  
 Db 78 GHEPEEGKYTFEERYDLVKFKVQVQAGLYVHLRIGPY 115  
 RESULT 2  
 YH20\_HAEIN STANDARD; PRT; 188 AA.  
 AC Q57066; O05085;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein H11720.  
 GN H11720.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OC NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kevlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudex D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Funtmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- SIMILARITY: BELONGS TO THE IS150/IS1296 ORF4 FAMILY.  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U32845; AAC23366.1; .  
 DR PIR; D64176; D64176.  
 DR TIGR; H11720; .  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 188 AA; 21747 MW; 3005CF9D4135F27 CRC64;  
 Query Match 23.0%; Score 64; DB 1; Length 188;  
 Best Local Similarity 32.4%; Pred. No. 0.87;  
 Matches 12; Conservative 10; Mismatches 15; Indels 0; Gaps 0;  
 Oy 8 GKRYNYRDSLVEAVKAVRGEMSVH-RAGSY 44  
 Db 75 GKRYNYRDSLVEAVKAVRGEMSVH-RAGSY 111  
 RESULT 3  
 TKR\_DROME STANDARD; PRT; 753 AA.  
 AC P14083;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Protein TKR.  
 GN TKR.  
 OS Drosophila melanogaster (fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88112827; PubMed=3428600.  
 RA Haller J., Cole S., Broenner G., Jaekle H.;  
 RT "Dorsal and neutral expression of a tyrosine kinase-related Drosophila  
 RT gene during embryonic development.";  
 RL Genes Dev. 1:862-867(1987).  
 CC -1- FUNCTION: POSSIBLE REGULATORY ROLE DURING DEVELOPMENT.  
 CC -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE A KINASE ON THE BASIS OF  
 CC WEAK AND NON-SIGNIFICANT SIMILARITIES.  
 CC PIR; A27041; A27041.  
 DR FLYBase; FBgn0003715; TKR.  
 DR Pfam; PF05225; HTH\_psq.1.  
 FT DOMAIN 143 151 POLY-ASP.  
 FT DOMAIN 153 157 POLY-GLU.  
 FT DOMAIN 174 183 POLY-ALA.  
 FT DOMAIN 221 224 POLY-ASN.  
 FT DOMAIN 297 306 POLY-ALA.  
 FT DOMAIN 325 332 POLY-ALA.  
 FT DOMAIN 709 712 POLY-ALA.  
 SQ SEQUENCE 753 AA; 81021 MW; F98D3272A7DDBE5E CRC64;  
 Query Match 23.0%; Score 64; DB 1; Length 753;  
 Best Local Similarity 28.6%; Pred. No. 3.8;  
 Matches 12; Conservative 13; Mismatches 17; Indels 0; Gaps 0;  
 Oy 5 PKRGYNYRDSLVEAVKAVRGEMSVH-RAGSY 46  
 Db 507 PKRGYNYRDSLVEAVKAVRGEMSVH-RAGSY 548  
 RESULT 4  
 FG4\_MOUSE STANDARD; PRT; 202 AA.  
 AC P11403; P15657;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibroblast growth factor-4 precursor (FGF-4) (K-fibroblast growth  
 DE factor) (HBGF-4).  
 GN FG4 OR FGF-4 OR KRGF.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89296455; PubMed=2740210;  
 RA Dickson C.;  
 RT "The mouse homologue of hsc/k-FGF: sequence, genome organization and  
 RT location relative to hnt-2.";  
 RL Nucleic Acids Res. 17:4037-4045(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90201563; PubMed=2318343;  
 RA Hebert J.M., Basilico C., Goldfarb M., Haub O., Martin G.R.;  
 RT "Isolation of cDNAs encoding four mouse FGF family members and  
 RT characterization of their expression patterns during embryogenesis.";  
 RL Dev. Biol. 138:454-463(1990).  
 CC -1- FUNCTION: IS ESSENTIAL FOR SURVIVAL OF THE POSTIMPLANTATION MOUSE  
 CC EMBRYO AND AT LATER EMBRYONIC STAGES, IS AN ESSENTIAL COMPONENT OF  
 CC SIGNALING NETWORK REQUIRED FOR GROWTH AND PATTERNING OF THE  
 CC DEVELOPING LIMB.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE BLASTOCYST INNER CELL MASS  
 CC AND LATER IN DISTINCT EMBRYONIC TISSUES.  
 CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X14849; CAA32967.1; -
CC EMBL: M30642; AAA37619.1; -
CC PIR: S04741; TVMSHS.
CC HSSP: P09038; 1BFG.
CC GO: MGI:95518; Fgf4.
CC GO: GO:0042475; P:odontogenesis (sensu Vertebrata); IDA.
CC InterPro: IPR002348; IL1_HBGF.
CC Pfam: PF00167; FGF_1.
CC PRINTS: PR00262; IL1HBGF.
CC ProDom: PD000831; IL1_HBGF; 1.
CC SMART: SM00442; FGF; 1.
CC PROSITE: PS00247; HBGF_FGF; 1.
CC Proto-oncogene: Growth_Factor; Mitogen; Signal.
CC SIGNAL 1 29
CC CHAIN 30 202
CC CONFLICT 167 167 A -> S (IN REF. 2).
CC FT 167 167
CC SQ SEQUENCE 202 AA; 21902 MW; 62D456231047CA31 CRC64;

Query Match 22.3%; Score 62; DB 1; Length 202;
Best Local Similarity 38.8%; Pred. No. 1.7;
Matches 19; Conservative 5; Mismatches 13; Indels 12; Gaps 2;

OY 15 RSLVAVKAVRGEMSV-----HRAQSYGVHSTLEYKKE 52
DB 108 RSLVLE-LSPVGRGVSVIFGVASRFVAMSSRGKLFGEVFFDECKFKE 155

RESULT 5
LEDB DAUCA STANDARD; PRT; 555 AA.
ID LED8 DAUCA
AC P20075;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Embryonic protein DC-8.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Queen Anne's Lace;
RX MEDLINE=89384429; PubMed=2571069;
RA Franz G., Hatzopoulos P., Jones T.J., Krauss M., Sung Z.R.;
RT "Molecular and genetic analysis of an embryonic gene, DC 8, from
RT Daucus carota L.";
RL Mol. Genet. 218:143-151(1989).
CC - FUNCTION: MAY PLAY A ROLE IN LATE EMBRYOGENY.
CC - SUBCELLULAR LOCATION: CYTOSOL; PROTEIN BODIES, AND CELL WALLS
CC - OF ZYGOTIC EMBRYO AND ENDOSPERM TISSUE.
CC - SIMILARITY: BELONGS TO THE LEA TYPE 1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16131; CAA34258.2; -
CC PIR: S04909; S04909.
CC InterPro: IPR004238; LEA.

```

```

DR Pfam: PF02987; LEA; 6.
KW Repeat.
FT DOMAIN 97 391 17 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 97 114 1.
FT REPEAT 115 125 2.
FT REPEAT 126 140 3.
FT REPEAT 141 154 4.
FT REPEAT 155 176 5.
FT REPEAT 177 191 6.
FT REPEAT 192 205 7.
FT REPEAT 206 216 8.
FT REPEAT 217 237 9.
FT REPEAT 238 259 10.
FT REPEAT 260 281 11.
FT REPEAT 282 303 12.
FT REPEAT 304 325 13.
FT REPEAT 326 343 14.
FT REPEAT 344 358 15.
FT REPEAT 359 376 16.
FT REPEAT 377 391 17.
SQ SEQUENCE 555 AA; 60260 MW; D15E8A30E51BD1AB CRC64;

Query Match 21.0%; Score 58.5; DB 1; Length 555;
Best Local Similarity 37.8%; Pred. No. 13;
Matches 14; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

OY 3 TRPKRGKRYNDRSLVLA-VKAVRGEMSVHRAQSY 38
DB 118 TWGKAGEYKDYTAQKAEEAKEKQAKAEETKAGEY 154

RESULT 6
MALQ MYCTU STANDARD; PRT; 724 AA.
ID MALQ MYCTU
AC O53932;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
DE (Disproportionating enzyme) (D-enzyme).
GN MALQ OR RV1781C OR MT1831 OR MTW049.03C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jorgels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultun J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Pleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Petersen J., Debby R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Ueberback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC - CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
CC to a new 4-position in an acceptor, which may be glucose or (1,4)-

```

alpha-D-glucan.  
 -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 -1- SIMILARITY: BELONGS TO THE DISPROPORTIONATING ENZYME FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: AL022021; CAAL7703.1; -  
 CC EMBL: AE007042; AAK46101.1; -  
 CC PIR: G70928; G70928.  
 CC TIGR: MT1831; -  
 CC TIGR: MT1831; -  
 CC Tuberculin; Rv1781c; -  
 CC InterPro: IPR003385; Glyco\_hydro\_77.  
 CC Pfam: PF02446; 4A\_glucohydrolase; 1.  
 CC TIGRFAMs: TIGR00217; ma0Q; 1.  
 CC Transferase; Glycosyltransferase; Carbohydrate metabolism;  
 CC Complete proteome.  
 CC SEQUENCE 724 AA; 79744 MW; 153B04525DE738EA CRC64;  
 -----  
 Query Match 21.0%; Score 58.5; DB 1; Length 724;  
 Best Local Similarity 30.8%; Pred. No. 17; Mismatches 20; Indels 7; Gaps 2;  
 Matches 16; Conservative 9;  
 -----  
 QY 2 GTRPKGRYRNDRLSLVEAVKAVQSGMSVHRAGS-YYGVPHSTLEKYKE 52.  
 Db 490 GAPFGQGYVYRVDHDMIGIV-----ALEAHRAAVAVVGGDLGVPEWVD 535  
 -----  
 RESULT 7  
 YJ59\_AQUAE STANDARD; PRT; 160 AA.  
 ID YJ59\_AQUAE  
 AC 067771;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein AQ\_1959 precursor.  
 GN AQ\_1959.  
 OS Aquifex aeolicus.  
 CC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 CC NCBI\_TaxID=63363;  
 CC RX MEDLINE=9819666; PubMed=9537320;  
 CC STRAIN=VF5;  
 CC Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 CC Graham D.E., Overbeek R., Snead M.A., Keller M., Anjey M., Huber R.,  
 CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 CC "The complete genome of the hyperthermophilic bacterium Aquifex  
 CC aeolicus";  
 CC Nature 392:353-358 (1998).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000765; AAC07742.1; -  
 CC PIR: G70467; G70467.  
 CC Hypothetical protein; Signal; Complete proteome.  
 CC SIGNAL 1 27  
 CC CHAIN 160  
 CC SEQUENCE 160 AA; 17700 MW; A3B5DB26BDBD548 CRC64;  
 -----  
 Query Match 20.7%; Score 57.5; DB 1; Length 160;  
 Best Local Similarity 29.1%; Pred. No. 4.7;

Matches 16; Conservative 9; Mismatches 9; Indels 21; Gaps 3;  
 -----  
 QY 8 KRYNYPD---RDSL-----VEAVKAVQSGMSVHRA-----GSYGV 41  
 Db 45 KRYNYPDGLRDIRPCKKVERVEVDPKGLKGLKGLKSLKVIASCNCAFQV 99  
 -----  
 RESULT 8  
 YJ51\_SHISO STANDARD; PRT; 100 AA.  
 ID YJ51\_SHISO  
 AC P16939;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Insertion element IS600 hypothetical 11 kDa protein (ISO-S3 11 kDa  
 DE protein).  
 OS Shigella sonnei.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Shigella.  
 CC NCBI\_TaxID=624;  
 CC RX MEDLINE=88062685; PubMed=2824781;  
 CC Matsutani S., Ohtsubo H., Maeda Y., Ohtsubo E.;  
 CC "Isolation and characterization of IS elements repeated in the  
 CC bacterial chromosome";  
 CC J. Mol. Biol. 196:445-455 (1987).  
 CC -1- SIMILARITY: BELONGS TO THE TRANSPOSASE FAMILY 8.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X05952; CA29384.1; -  
 CC PIR: S03411; S03411.  
 CC InterPro: IPR002514; Transposase\_8.  
 CC Pfam: PF01527; Transposase 8; 1.  
 CC KW Hypothetical protein; Transposase element.  
 CC SEQUENCE 100 AA; 11664 MW; 6C32F3659F1B361A CRC64;  
 -----  
 Query Match 20.5%; Score 57; DB 1; Length 100;  
 Best Local Similarity 28.9%; Pred. No. 3.3;  
 Matches 11; Conservative 8; Mismatches 19; Indels 0; Gaps 0;  
 -----  
 QY 9 KRYNYPDRLSLVEAVKAVQSGMSVHRAGSYGVPHSTL 46  
 Db 4 KTORYSKEFKRAVRYTPENDLSISGASRLSLPEGLT 41  
 -----  
 RESULT 9  
 YJUA\_SCHPO STANDARD; PRT; 128 AA.  
 ID YJUA\_SCHPO  
 AC O9C1W5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C713.10 in chromosome II.  
 GN SPBC713.10.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 CC NCBI\_TaxID=4896;  
 CC RX MEDLINE=21848401; PubMed=11859360;  
 CC Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 CC Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Felwell J., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jasele K.,  
 RA James K., Jones L., Jones M., Leach S., McDonald S., McLean U.,  
 RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbitts E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grynolprez B.,  
 RA Welter J., Vanterre E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Mambur R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas R., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakowski G.V., Useery D., Barrell B.G., Nurse P.,  
 RA "The genome sequence of Schizosaccharomyces pombe".  
 RL Nature 415:871-880(2002).  
 CC -I- SIMILARITY: BELONGS TO THE UPF0108 (MAGMAS) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL512943; CAC22611.1; -  
 DR "GeneDB: SPombe; SPBC713.10; -  
 DR InterPro: IPR005341; UPF0108.  
 DR Pfam: PF03656; UPF0108; 1.  
 DR ProDom: PD311402; UPF0108; 1.  
 KW Hypothetical protein.  
 KW SEQUENCE 128 AA; 14120 MW; 912CB3B6B5A58FC CRC64;  
 SO  
 Query Match 20.3%; Score 56.5; DB 1; Length 128;  
 Best Local Similarity 41.9%; Pred. No. 4.9;  
 Matches 13; Conservative 8; Mismatches 9; Indels 1; Gaps 1;  
 Oy 24 AVORGEMSVHRAGSYGV-PHSTLEYKYER 53  
 Db 49 AVRREGMTIOEGSIINIKPESLEGELEKR 79  
 RESULT 10  
 NIFK\_RHISN STANDARD; PRT; 513 AA.  
 ID NIFK\_RHISN  
 AC P13067;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Nitrogenase molybdenum-iron protein beta chain (EC 1.18.6.1)  
 DE (Nitrogenase component I) (Dinitrogenase)  
 GN (NIFK1 OR Y4VM) AND (NIFK2 OR Y4XC).  
 OS Rhizobium sp. (strain NGR234).  
 CC Plasmid BYM pNGR234.  
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9730556; PubMed=9163424;  
 RA Freiberg C.A., Fellay R., Baitoch A., Broughton W.J., Rosenthal A.,  
 RA Perret X.,  
 RT "Molecular basis of symbiosis between Rhizobium and legumes".  
 RL Nature 387:394-401(1997).

RN [2]  
 CC SEQUENCE OF 132-195 FROM N.A.  
 RP STRAIN-ANU 240.  
 RC MEDLINE=8930671; PubMed=2744485.  
 RX Badenoch-Jones J., Holton T.A., Morrison C.M., Scott K.F., Shine J.,  
 RA "Structural and functional analysis of nitrogenase genes from the  
 RT broad-host-range Rhizobium strain ANU240."  
 RL Gene 77:141-153(1989).  
 CC -I- FUNCTION: THE KEY ENZYMAIC REACTIONS IN NITROGEN FIXATION ARE  
 CC CATALYZED BY THE NITROGENASE COMPLEX, WHICH HAS 2 COMPONENTS: THE  
 CC IRON PROTEIN AND THE MOLYBDENUM-IRON PROTEIN.  
 CC -I- CATALYTIC ACTIVITY: 8 reduced ferredoxin + 8 H(+) + N(2) + 16 ATP  
 CC = 8 oxidized ferredoxin + 2 NH(3) + 16 ADP + 16 phosphate.  
 CC -I- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS THAT BINDS  
 CC 30-32 FE, 2 MO, AND INORGANIC SULFUR.  
 CC -I- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M26963; AAA26327.1; -  
 DR EMBL: AE000102; AAB91901.1; -  
 DR EMBL: AE000105; AAB91925.1; -  
 DR PIR: PS0046; PS0046.  
 DR PIR: T10830; T10830.  
 DR HSSP: P07329; 3MIN.  
 DR InterPro: IPR005976; NIFK.  
 DR InterPro: IPR000318; Nitrogenase\_comp1.  
 DR InterPro: IPR000510; Oxidored\_nitrogenase.  
 DR Pfam: PF00148; oxidored\_nitro; 1.  
 DR TIGRfam: TIGR01286; nifK; 1.  
 DR PROSITE: PS00699; NITROGENASE\_1; 1.  
 DR PROSITE: PS00900; NITROGENASE\_1\_2; 1.  
 KW Oxidoreductase; Nitrogen fixation; Molybdenum; Iron-sulfur; Plasmid;  
 KW Multigene family.  
 KW SEQUENCE 513 AA; 57302 MW; 41631040335541AE CRC64;  
 SO  
 Query Match 20.3%; Score 56.5; DB 1; Length 513;  
 Best Local Similarity 32.6%; Pred. No. 21;  
 Matches 15; Conservative 8; Mismatches 16; Indels 7; Gaps 2;  
 Oy 5 PKRGKRYNRDRLVAVKAVORGE--MSVHRAGSYGVPHSTLEY 48  
 Db 262 PSDGEYRMVDOGTTIKALRALNMEATLSLGHVNS-----RKTLEY 302  
 RESULT 11  
 FGFM\_HUMAN STANDARD; PRT; 170 AA.  
 ID FGFM\_HUMAN  
 AC O9HCT0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fibroblast growth factor-22 precursor (FGF-22).  
 GN FGF22.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP TISSUE=Placenta;  
 RX MEDLINE=21240339; PubMed=11342227;  
 RA Nakatani Y., Hosonaka M., Asaki T., Kaesai Y., Itoh N.,  
 RA "Identification of a novel fibroblast growth factor, FGF-22,  
 RT preferentially expressed in the inner root sheath of the hair  
 RT follicle".  
 RL Biochim. Biophys. Acta 1517:460-463(2001).

CC -1- FUNCTION: MAY BE INVOLVED IN HAIR DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB021925; BAB13479.1; -  
 DR HSSP: P31371; 1682.  
 DR Genew: HGNC:3679; FGF22.  
 DR MIM: 605831; -  
 DR GO: GO:0005615; C:extracellular space; NAS.  
 DR GO: GO:0030154; P:cell differentiation; NAS.  
 DR InterPro: IPR002348; IL1\_HBGF.  
 DR Pfam: PF00167; FGF\_1.  
 DR PRINTS: PR00262; IL1HBGF.  
 DR ProDom: PD000831; IL1\_HBGF; 1.  
 DR SMART: SM00442; FGF\_1.  
 DR PROSITE: PS00247; HBGF\_FGF; FALSE\_NEG.  
 DR Growth factor; Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 170 FIBROBLAST GROWTH FACTOR-22.  
 SQ SEQUENCE 170 AA; 19662 MW; CB88918CD54ACE; CRC64;  
 Query Match 20.1%; Score 56; DB 1; Length 170;  
 Best Local Similarity 28.6%; Pred. No. 7.6;  
 Matches 18; Conservative 14; Mismatches 15; Indels 16; Gaps 3;  
 QY 1 KGTPEKKGKYNVDRDLVE-----AYKAVQRC-EMVVRAGSTGYVPHSTLEKV 50  
 DB 62 QGTWRHKG-----QDSILIRSVHVGVVVKAIVSSGFGVAMNRGRGLYVDCRF 115  
 QY 51 KER 53  
 DB 116 RER 118  
 RESULT 12  
 BAB1 DROME STANDARD: PRT; 977 AA.  
 AC Q9W0K7; Q23968; Q8WR78; Q9U1H7;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DE Bric-a-brac protein 1.  
 DE BAB1 OR BAB OR CG9097/CG13910.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pezomyzeta;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM B), FUNCTION, SUBCELLULAR LOCATION, AND  
 RP TISSUE SPECIFICITY.  
 RC TISSUE=Larva, and Ovary;  
 RX MEDLINE=21969340; PubMed=11973274;  
 RA Couderc J.L.G., Godt D., Zollman S., Chen J., Li M., Tjong S.,  
 RA Cramton S.E., Sabut-Barnola I., Laski F.A.;  
 RA "The bric a brac locus consists of two paralogous genes encoding  
 RA BTB/POZ domain proteins and acts as a homeotic and morphogenetic  
 RA regulator of imaginal development in Drosophila.";  
 RL Development 129:2419-2433(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galile R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abbill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baau A., Bakken D., Andrews-Pfannkoch L., Basley E.M.,  
 RA Beeson K.Y., Berens P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Bokoyva D., Borhan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP REVISIONS.  
 RC STRAIN=Berkley;  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Dysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RA systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM A).  
 RC STRAIN=Berkley; TISSUE=Testis;  
 RX MEDLINE=22426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celniker S.E.;  
 RA "A Drosophila full-length cDNA resource.";  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 RN [5]  
 RP SEQUENCE OF 99-225 FROM N.A.  
 RX MEDLINE=95280944; PubMed=7760839;  
 RA Chen W., Zollman S., Couderc J.L., Laski F.A.;  
 RA "The BTB domain of bric a brac mediates dimerization in vitro.";  
 RL Mol. Cell. Biol. 15:3424-3429(1995).  
 CC -1- FUNCTION: Probably acts as a transcriptional regulator. Required  
 CC for the specification of the tarsal segment. Also involved in  
 CC antenna development.  
 CC -1- SUBUNIT: May form dimers.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=B;

```
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditiidae; Peloiderinae; Caenorhabditis.  
OX NCBI_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ahnescoug R., Anderson K., Baynes C., Berks M.,  
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
Coxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
Johnson L., Jones M., Kershaw J., Kirken J., Laister N.,  
Larrelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,  
Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,  
Sturison J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,  
Woldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans." ;  
RL Nature 368:32-38(1994).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/  
or send an email to licenc@esb-gib.ch).  
-----  
DR EMBL, LO7144; AKK2144.1; .  
DR PIR, S44862; S44862.  
DR Wormpep; R0SD3_2; CE00281.  
DR InterPro; IPR006876; LMBR1.  
DR Pfam; PF04791; LMBR1, 1.  
KM Hypothetical protein.  
SQ SEQUENCE 737 AA; 83555 MW; 3397543C5CECB9B4 CRC64;  
  
Query Match 20.0%; Score 55.5; DB 1; Length 737;  
Best Local Similarity 35.9%; Pred. No. 42;  
Matches 14; Conservative 7; Mismatches 17; Indels 1; Gaps 1;  
  
QY 5 PKRGKYRNYRDLSLEAVKAVGSGSVHRSAGSY-YGVV 42  
|||:::||::|:|||||::|:|:  
Db 407 PKRKPNENFDYRNLIKVEYKEARRQRSSLDSEDDVFEGSP 445  
  
RESULT 14  
DEGU_BACBR STANDARD; PRT; 236 AA.  
ID DEGU_BACBR STDIN; PRT; 236 AA.  
AC PS4662;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transcriptional regulatory protein degU.  
GN DEG U.  
OS Bacillus brevis (Brevibacillus brevis).  
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.  
OX NCBI_TaxID=1393;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ALK36;  
RX MEDLINE=95169370; PubMed=7765823;  
RA Louw M.E., Reid S.J., James M.D., Watson T.G.;  
RT Cloning and sequencing the degS-degu operon from an alkalophillic  
RT Bacillus brevis."; RL Appl. Microbiol. Biotechnol. 42:78-84(1994).  
-- FUNCTION: REGULATING FACTOR FOR THE PRODUCTION OF EXTRACELLULAR  
CC PROTASES. THE N-TERMINAL REGION ACTS AS AN INHIBITOR, WHEREAS  
CC THE C-TERMINAL REGION CARRIES ENHANCING ACTIVITY.  
CC -- SIMILARITY: Contains 1 response regulatory domain.  
CC -- SIMILARITY: BELONGS TO THE LuxR/UHPA FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L15444; AAC41439.1; -
DR PIR; J39835; I39835.
DR HSSP; P10957; IRLN.
DR InterPro; IPR000792; HTH LuxR.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00196; Gere; 1.
DR Pfam; PF00072; response_reg; 1.
DR PRINTS; PR00038; HTH LuxR.
DR ProDom; PD000307; HTH LuxR; 1.
DR ProDom; PD000309; Response_reg; 1.
DR SMART; SM00421; HTH LuxR; 1.
DR SMART; SM00448; REC_1.
DR PROSITE; PS00622; HTH LuxR FAMILY; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR Sensory transduction; Phosphorylation; Transcription regulation;
KM DNA-binding; Activator; Repressor.
FT DOMAIN 11 127 RESPONSE REGULATORY.
FT MOD_RES 62 62 PHOSPHORYLATION (BY SIMILARITY).
FT DNAS_BIND 190 209 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 236 AA; 27003 MW; 9E464265B0365315 CPC64;

Query Match 19.8%; Score 55; DB 1; Length 236;
Best Local Similarity 35.0%; Pred. NO. 14;
Matches 14; Conservative 8; Mismatches 12; Indels 6; Gaps 2;

Qy 11 RNYRDSLVEAVKAVQGEVSRNAGSYGVPHSTL-EYK 49
Db 112 KEMDADALIEAVKVAQCAVYHPR-----VTHNLIKEYR 146

RESULT 15
CST1 HUMAN STANDARD; PRT; 431 AA.
AC Q005048;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cleavage stimulation factor, 50 kDa subunit (CSTF 50 kDa subunit)
DE (CF-1 50 kDa subunit).
GN CSTF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RX SEQUENCE FROM N.A., AND SEQUENCE OF 101-119 AND 155-170.
RP MEDLINE=93054692; PubMed=1358884;
RA Takagaki Y., Manley J.L.;
RT "A human polyadenylation factor is a G protein beta-subunit
RT homologue."
RL J. Biol. Chem. 267:23471-23474(1992).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blake S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Bridgman A.M., Brown A.J.,
RA Chapman J.C., Clamp M., Collier R.E., Connor R.E., Corby N.R.,
RA Clegg S., Copley V.E., Collier R.E., Clark L.N., Clark S.Y., Clee C.M.,
RA Coulson A., Coville G.U., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

```

```

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levesialho M.H., Levesia M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McDonachie L.J., McMay K., McMurtry A.A.,
RA Mine S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whithead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871(2001).
RN 13
RP SEQUENCE FROM N.A.
RX TISSUE=Muscle;
RX PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: ONE OF THE MULTIPLE FACTORS REQUIRED FOR POLYADENYLATION
CC AND 3'-END CLEAVAGE OF MAMMALIAN PRE-MRNAs. MAY BE RESPONSIBLE
CC FOR THE INTERACTION OF CSTF WITH OTHER FACTORS TO FORM A STABLE
CC COMPLEX ON THE PRE-mRNA.
CC -1- SUBUNIT: COMPOSED OF THREE DISTINCT SUBUNITS OF 77, 64, AND 50
CC kDa.
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: Contains 6 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L025947; AAA35691.1; -
DR EMBL; AL121914; CAC12718.1; -
DR EMBL; BC001011; AAH01011.1; -
DR EMBL; BC007425; AAH07425.1; -
DR PIR; A45142; A45142.
DR Genew; HGNC:2483; CSTF1.
DR MIM; 600369; -
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003723; F:RNA binding activity; TAS.
DR GO; GO:0006379; P:mRNA cleavage; TAS.
DR GO; GO:0006378; P:mRNA polyadenylation; TAS.
DR GO; GO:0006396; P:RNA processing; TAS.
DR InterPro; IPR001680; WD40.

```

DR Pfam; PF00400; WD40; 6.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR ProDom; PD000018; WD40; 1.  
DR SMART; SM00320; WD40; 6.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
DR PROSITE; PS0082; WD\_REPEATS\_2; 4.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat; Nuclear protein.  
FT DOMAIN 14 35 HYDROPHOBIC.  
FT REPEAT 106 145 WD 1.  
FT REPEAT 171 210 WD 2.  
FT REPEAT 215 254 WD 3.  
FT REPEAT 260 301 WD 4.  
FT REPEAT 303 343 WD 5.  
FT REPEAT 395 430 WD 6.  
SQ SEQUENCE 431 AA; 48357 MW; 88A5B53022AD9E3 CRC64;

Query Match 19.8%; Score 55; DB 1; Length 431;  
Best Local Similarity 33.3%; Pred. No. 27;  
Matches 17; Conservative 6; Mismatches 22; Indels 6; Gaps 2;

QY 2 GTRPKRGKRYNYDRDGLVEAVKAVORGEN---SVHRAGSY--YGVPHSTL 46  
Db 191 GSRDYTLKLFDPYKSPSAKRAFKYIOEAEMLRISIFHPGDFILVGTQHPTL 241

Search completed: October 28, 2003, 12:02:33  
Job time : 5.92727 secs

**THIS PAGE BLANK (USPTO)**



A:Residues: 1-1109 <TAN>  
A:Cross-references: GB:MS109; NID:gl65692; PIDN:AAA3162.1; PID:gl65693  
R:Ident, P.; Campbell, D.G.; Hubbard, M.J.; Cohen, P.  
FEBS Lett. 248, 67-72, 1989  
A:Title: Multisite phosphorylation of the glycogen-binding subunit of protein phosphatase  
A:Reference number: S04004; MUID:89252053; PMID:2542090  
A:Accession: S04004  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 33-68 <DEN>  
C:Keywords: phosphoprotein; phosphoric monoester hydrolase

Query Match 5.9% Score 133.5; DB 2; Length 1109;  
Best Local Similarity 22.9%; Pred. No. 1.4; Mismatches 163; Indels 83; Gaps 17;  
Matches 89; Conservative 53; Mismatches 163; Indels 83; Gaps 17;

7 QFAIEYIKSGKTQENRNGSIGPSIVCKSIQMNQANSLQEOEGPDLITVNRMQEONTQ 66  
Db EFCEIRYETVSGTFPMNNNGT-NYTLVQ-----KKEPEPEPKPL-----EAPSK 252  
67 QGDGVLDTKTKTSIKSEESSICDPSSSENSVAGRLHRRREDYVER--SAEFADGLIS-- 121  
Db QKGGCLKYKSSK-----EESS--ETSENNFEN--SKADYPIPIYVCSHEEKEDLKSSY 303  
122 KALKIOGSDALINKAGLYGIPQKTLHLHLEALPAGK-----PASPKNTRDFHDSYS 175  
Db QNVKQVNTDEHNEKELELMINORLITRCAASYGKNTLSDPSNIPNPEELQKKQS 363  
176 YKDSKETCAVLQKVALMARAQERTREKSKLNLLETSEIKFPASTYHLQTLQKMTQFK 235  
Db H---SEACTDLSQRLSLPGSSAESLSLKGDFHTE---KYSSGNSSHQ-----PDMG 410  
236 EKNSLOYETSNPTVQLKIPOLRVSSVSKSPDPDGSGLDV-----MYQVSKTSSVLEGS 289  
Db EINPSLGGTTSDGSVQLHSSKEILDDNANPAHGSGRGEICSPFGQLKASLNKKYEGG 470  
290 ALQKLKNIPLPQNKIKCEGPTVTHSSVDSYFLHGDLSPLCLNSKNGTVDTGSENTDGLDR 349  
Db AENS-----EMKQCECLPRDVHLKASDYF-----KSTENRPSSE--EDYGTSS 510  
350 KDSKQPR-----KKRGRYQYDHE 368  
Db KDNKEKRIQLDVEDETSKNFRSIFPDQE 538

RESULT 3  
T43523  
cut17 protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Jun-2000  
C:Accession: T43523; T41649; T41700  
R:Morishita, J.; Matsusaka, T.; Yanagida, M.  
Submitted to the EMBL Data Library, August 1999  
A:Description: Fission yeast cut17 is required for chromosome segregation.  
A:Reference number: Z22536  
A:Accession: T43523  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-997 <MOR>  
A:Cross-references: EMBL:AB031034; PIDN:BAAB83415.1  
R:Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
Submitted to the EMBL Data Library, August 1998  
A:Reference number: Z22007  
A:Accession: T41649  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-997 <HAR>  
A:Cross-references: EMBL:AL031333; PIDN:CAA20434.1; GSPDB:GN00068; SPDB:SPCC962.02C  
A:Experimental source: strain 972h-; cosmid c962  
R:Wiedler, H.; Duessehoeft, A.; McPougal, R.C.; Rajandream, M.A.; Barrell, B.G.  
Submitted to the EMBL Data Library, October 1999  
A:Reference number: Z22010  
A:Accession: T41700

[illegible]

```

Db 273 GCHDPSVISPSTDMKNIVKALQNFQD-----IANTLNTVSSAQSGGI-DISLYKQIYKMK 326
Qy 85 ESSICDPSSSENSVAGRLNRNREDYVERSAEPADGLSLKALKDIOGALDINKAGILYGP 144
Db 327 YDFVEDPNGKYSV-----DKDKF-----DKLYKALMGFETNLAG-EYGI- 366
Qy 145 OKTLILAH-EALP-----AGKPASFQKTRDPFH-----DSYKDSME 181
Db 367 -KTRSYSEVLPPIKTEKLDNTTYTONEGNISKULKTEFNQONKAANKAEAEISL 425
Qy 182 TCVALQKVALMARAOAERTESKLNLETSEIKFPASTAYLHQLTQKWTQFKENESL 241
Db 426 EHLVYRIAMCKPVMYKATGKSECCIVNNEDLFFIAN-----KDSFKDLAKAETI 477
Qy 242 QYETSNPTVQ-----LKIPI-QLRVSSVSKOP 267
Db 478 AYNTONNTNTEENNFSIDOLILNDLSSGIDLPNENTEPTNPDIDIPYVIXOSALKIPV 537
Qy 268 DSGGLDVMYQVSKTSSVLEGSALQKXNLPKONK-----IEGSPVTHSSVDSY 318
Db 538 DGDLSFEVLHAOTFPNSNIENLQLTNSLNDALNNKVTTFESTNLVERKANTVVGAS---- 593
Qy 319 FLHGDSLPLCLNSKNGTVDG-TSENTE 344
Db 594 -----LPVWVWKGVIDDFTSESTQ 612

```

## RESULT 5

```

119172
hypochemical protein F18C12.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
R/Harris, B.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19083
A/Accession: T19172
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-545 <W1>
A/Cross-references: EMBL:Z81466; PIDN:CA803870.1; GSPDB:GN00019; CESP:F18C12.3
A/Experimental source: clone C09H6
R/Harris, B.
submitted to the EMBL Data Library, June 1996
A/Reference number: Z19371
A/Accession: T21088
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-545 <W1>
A/Cross-references: EMBL:Z75536; PIDN:CA899833.1; GSPDB:GN00019; CESP:F18C12.3
A/Experimental source: clone F18C12
A/Map position: 1
A/Introns: 171/3; 222/2; 316/3; 368/3; 409/1; 409/3; 493/1

```

```

Query Match 5.8%; Score 129.5; DB 2; Length 545;
Best Local Similarity 20.8%; Pred. No. 0.96;
Matches 93; Conservative 72; Mismatches 174; Indels 109; Gaps 20;
Qy 2 KKMIRQFA-----IEYISKSGKTOENNGSIGPISVCKSIOMNOAENSLOE---EQEGP 52
Db 175 KDSVAKFPALNTKSLAEIKNGKTEVAER--LMYLVQOSIIVKIQ-SIREKYTEWNT 231
Qy 53 LDLTVMRQEQNTQOGDGLDLSTKKTSTI--KSESSICDPSSSENSVAGRLHRRREYV 109
Db 232 AAFNQRIGQGOYANKPEKSKSTKSTKELINENLEDNNEDEDSNEKEEIEENBEDYD 291
Qy 110 ERSAAEFADGLSKALKDIOGALDINKAGILYGIPOKTLHLLEAL-----PACKPASF 163
Db 292 QSDIEMLDS-----DEEAGGEAAKARRNLLGLIGVQEDKRPPLAPKRR 337

```

```

Qy 164 KNKTRDFHDSYSYKDSKETCAVLQKVALMARAOAERTESKLNLETSEIKFPASTAYLH 223
Db 338 KLVEEDEVETVAKQKMKROIPEKE-----LKSVEVKNP----- 372
Qy 224 QLTLOKWTQFKENESIQYETSNPTVOLKIPQLRV--SVYSKQPDGSGGLDVMYQVSK 281
Db 373 -----KSKNSTKPKPTKSAPIVKIKVEEKEVEENVEDSDQKTLV-MKVDLSK 420
Qy 282 TSSVLEGSALQKXNLPKONKIECSGPVTHSSVD-----SYFL--HDDLSPCLNSKNG 334
Db 421 GGIKAKA---QKFTTAPKSAKI--VAPVSEDENDDESSFFLPKSGVAPRKTIIPK- 474
Qy 335 TVDGTSEVTEEDGLPKDSKOPKKRGRYQVDEIMEBAIMWNSGKMSVKAQGIYGP 394
Db 475 -----PSEVKEK-VDKRFPKGQYK-----SEAVVEKKKG-----SKSAVSGEM 513
Qy 395 H-STLEYVKERSGTLTKPPKKRLPLP 421
Db 514 HPSWIASQLKKKELASAKPCGKKITFGD 541

```

## RESULT 6

```

142727
proliferation potential-related protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
R/Witte, M.M.; Scott, R.E.
A/Accession: T42727
A/Reference number: Z22246
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1560 <W1>
A/Cross-references: EMBL:U83913; NID:G3858884; PID:G3858885; PIDN:AACT2432.1
A/Experimental source: strain Balb/C
A/Genetics:
A/Genes: P2P-R
A/Function:
A/Description: involved in hnRNP association and Rb1 binding
C/Superfamily: RING finger homology
F:57-107/Domain: RING finger homology <RNV>

```

```

Query Match 5.8%; Score 129.5; DB 2; Length 1560;
Best Local Similarity 20.5%; Pred. No. 4;
Matches 103; Conservative 81; Mismatches 180; Indels 139; Gaps 23;
Qy 15 KSGKTOENRNGSIGPISVCKSIOMNOAENSLOEBOGPLDT-----VNMROEQNTQOG 68
Db 786 KSDTKRKS DG-----SATAKQDNVLKPSKGOEKYDGDREKSPSEPPKAKAKEATK-- 839
Qy 69 DGVLDLSTKKTSTINSESSICDPSSSENSVAGRLHNRN--REDYVERSAEPADGLSLKAL 124
Db 840 ---IDSVKPSSSSQDEKVTGTPRKAKSKAKHQEAKPADEKVKKCC-----SKDI 889
Qy 125 KDIOGALDINKAGILYGIPOKTLHLLEALPAKRPASFQKTRDFHDSYSYKDSKETCA 184
Db 890 KSEKPSAKD-EKA-----KKPEKNRLDSKGEKRRKTEKEKSVDPKF-ESSSMKISKVEGT 943
Qy 185 VLOKVALMARAOA-----ERT-EKSKLNLETSEIKFPASTAYLHQLTQKWTQFKENESL 241
Db 944 EIVVPSPRKMGEGVEKLEKTPKOKTASSTT-----PAKIKILRETGKIKGNENAST 998
Qy 234 FKKNESIQYETSNPTVOLKIPQLRVSSVSKQPDGSGGLDVMYQVSKTSSVLEGSALQ 292
Db 999 TKPESEKLESTSS-----KIKQEKVKGAKAKRYAVAGSGSSSTLVYDVTSTT-GGSPVR 1051
Qy 293 KTK-----NLP-----KQNKIECSGPVT 311
Db 1052 KSEKTDTRTVIKTMEBYNNDNTAPADVIMIHVPOSKWDKDFESEEDVDVKTQPIQ 1111
Qy 312 HSSVDSYFLHGDSLPLCLNSKNGTV--DGTSEVTEEDGLDRKDSKOPKKRGRYQVDEI 369

```

Db 1112 SVGRSSII-----KNVTTKPSATAKYTE-----KESPOPEKTLQKLPKASHL 1155  
 Qy 370 MEAIAVAMSGKSVSKAQGIYGVPHSLTEYKVERSGTLTPPKKL----- 417  
 Db 1156 MOHEI---RSSKGSASHSEK-----RAKDRHSSEKDNPKKSGAPDPESTVD 1203  
 Qy 418 RLPTGLYNTMTDSTGSCSKNSK 440  
 Db 1204 RLSEGHFKTLTSSSKERTSEK 1226

## RESULT 7

S61185

hypothetical protein YDR299w - Yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9740.7

C:Species: Saccharomyces cerevisiae

C&gt;Date: 23-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 19-Apr-2002

C/Accession: S61185

R: Ding, H.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of S. cerevisiae cosmid 9740.

A:Reference number: S61160

A/Accession: S61185

A:Molecule type: DNA

A:Residues: 1-534 &lt;DIN&gt;

A/Cross-references: EMBL:U28374; NID:g849207; PID:g849214; GSPDB:GN00004; MIPS:YDR299w

C:Genetics:

A:Gene: SCD-BPR2; MIPS:YDR299w

A/Cross-references: SCD:S0002707

A/Map position: 4R

Query Match 5.6%; Score 127; DB 2; Length 534;  
 Best Local Similarity 21.4%; Pred. No. 1.4;  
 Matches 92; Conservative 61; Mismatches 157; Indels 120; Gaps 16;

Qy 5 IRQPAIEYISKSGTQENRNGSI-----GPSIVCKSIQNNQAEINSIQEPEGPLDITV 57  
 Db 9 ISDIAIKPVNKDFIDEENASLFOHNEKNGES-----DLSDYGNSTETETKKAHYLEV 62  
 Qy 58 NRMQEQNGQGDVLDLSTK-TSISKRESSICDPSSENSVAGRLHRRREDYVERSAEF- 115  
 Db 63 ---EKSRLRAEKGLELNDPRTYGVKSGRQALYEVSSENEDEEEEEEKEEDALSFR 118  
 Qy 116 -----ADG-----LISKAL-----KDIQSGALDINKAGI 139  
 Db 119 TDSDEVEIDEESDADGGETEQAQKHALSKLIQETKQALINKLSQVQORDASKG-- 176  
 Qy 140 LYGIPQKTL-----LHLALPAGKPASFNKTRDFHDSYSYKDSKETCAVLQKVALMA 193  
 Db 177 -YSIIQQTCLFDNIIDLRILQKAVIAANKLPLTTESWEAKMDSEETKRLK----- 229  
 Qy 194 RAQAEKRTSKLNLLETSEIKF-----PTASTYLHQLTLQKAVTQKKNESIQYETS 246  
 Db 230 --ENEKLFNNLFNRIINFRIFQGDHITONEEVAKHLKSKRSIKELQYETNSIDSELK 287  
 Qy 247 N-PTVOLKIPQLRVSVSVSKQPDGSGL-----LPMVQVSKTSVLEGSALQKUKN 296  
 Db 288 EYRTAVLNKMTKTVSSASGNALSNKRAKALNPADVQVENVQLSMSMLMKRTKLN-RN 346  
 Qy 297 ILPKONKIEG-----GPVTHSSVDSYFLHGDLSPLCLNSKNGTVDGISENTEDGLD 348  
 Db 347 ITPLYFQKDCANGRLPELISPVKDSVDD-----NENSDGDL 384  
 Qy 349 RKDSKQPRKK 358  
 Db 385 IPKNDPFRKK 394

## RESULT 8

B49284

Immediate-early protein RF3/RF4 - human herpesvirus 6 (strain Z29) (fragment)

C:Species: human herpesvirus 6

C&gt;Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 08-Oct-1999

C/Accession: B49284  
 R:Chou, S.; Marousek, G.I.  
 Virolgy 198; 370-376, 1994  
 A:Title: Analysis of interstrain variation in a putative immediate-early region of human  
 A:Reference number: B49284; MUID:94082474; PMID:8259673  
 A/Accession: B49284  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-983 <CHO>  
 A/Cross-references: GB:I21760; NID:g347260; PID:AAA15547.1; PID:g347261

## Query Match

5.6%; Score 126.5; DB 2; Length 983;  
 Best Local Similarity 18.7%; Pred. No. 3.3;  
 Matches 73; Conservative 73; Mismatches 164; Indels 81; Gaps 13;

Qy 9 AIEYISKSGK-TQENRNGSIGPSIVCKSIQNNQAEINSIQEPEGPLDITVNRMOEQNTQ 67  
 Db 642 AVSQCKSKKRTAKKVPKPS-KSKKIKLDRPET-----TNVIVISSEDEED 692  
 Qy 68 GDGVLDSLTKKTSIKSESSICDPSSENSVAGRLHRRREDYVERSAEFADGLSKALD 127  
 Db 693 GNNIIDSMLKTIKSE-----PNSSESSSDCTSEDVYLH-----LSDYKVI 737  
 Qy 128 QSGALDINKAGILYGIPOKTLHLLEALPAGKPASFNKTRDFHDSYSYKDSKETCAVLQ 187  
 Db 738 NNGHCQSK-----GFPSVFITPIRSMPC-----THDIRNKF-----VPK 772  
 Qy 188 KVALMAAQAEKTE-----KSKNLLETSEIKFPTASTYLHQLTLQKAVTQKKNESL 241  
 Db 773 KHWLWFMKTHKVDNCVHSSAKKNVNDSDVTEAHNCFINHPVPIKTDDEEKEKENVSY 832  
 Qy 242 QY-----ETSNPTVOLKIPQLRVSVSVSKQPDGSGLDPVQVSKTSVLEGSAL 291  
 Db 833 TYSKIEDSKTDLIEDITTKLITEMWENMDLIDIKHGIARHCQDLSKYVITHTAC 892  
 Qy 292 QKLNILPKONKIEGSPVTHSSVDSYFLHGDLSPL-----CLNSKNGTVDGISENT 343  
 Db 893 EKNLNVANSQNLVTAETQIFDPQGT-----GNNSPLINIINDTTCQDENRECTGTSNDN 947  
 Qy 344 EDGLDRKDSKQPRKKRRTQY--DHEIMEE 372  
 Db 948 EKTIRSDCNSDKMEVFKLGDYSDYDFEE 978

## RESULT 9

T44232

hypothetical protein U90 [imported] - human herpesvirus 6 (strain Z29)

C:Species: human herpesvirus 6

A:Variety: strain Z29

C&gt;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Jun-2000

C/Accession: T44232

R:Dominguez, G.; Dambaugh, T.R.; Stamey, F.R.; Dewhurst, S.; Inoue, N.; Pellett, P.E.

J. Virol. 73; 8040-8052, 1999

A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with human

A:Reference number: 222734; MUID:99412318; PMID:10482553

A/Accession: T44232

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1078 &lt;DOM&gt;

A/Cross-references: EMBL:AF157706; PIDN:AA49675.1

A:Experimental source: strain Z29; variant B

C:Genetics:

A:Introns: 32/2; 103/1

A&gt;Note: U90

## Query Match

5.6%; Score 126.5; DB 2; Length 1078;  
 Best Local Similarity 18.7%; Pred. No. 3.7;  
 Matches 73; Conservative 73; Mismatches 164; Indels 81; Gaps 13;

Qy 9 AIEYISKSGK-TQENRNGSIGPSIVCKSIQNNQAEINSIQEPEGPLDITVNRMOEQNTQ 67  
 Db 737 AVSQCKSKKRTAKKVPKPS-KSKKIKLDRPET-----TNVIVISSEDEED 787

```

QY 68 GDGVLDLSTKKTJSIKSEESSICDPSESSENVAGRLHNHRDIVERSAEFPDGLLSKALXOI 127
Db 788 GNNITDKSMLEKTIYSE-----PNSSESSSEDDCTSEBNYLH-----LSDYDKVI 832
QY 128 QSGALDINKAGILYGIPOKTLILHLLEALPAGRPASFKNTRDPFHDSYSYKDSKETAVALQ 187
Db 833 NNGHQOSK-----GFPSFVPTIPIRSMG-----THDIRNKF-----VVK 867
QY 188 KVALIMARAQAERTE-----KSKNLLETSEIKFPASTYLLQTLQKXVTOFEKXNEBL 241
Db 868 KHWLWFMKRTKHVNDXCVIHSSAKMAYKNDSDVTEANHCFINFPVPIKTDDEYEKXNVSQY 927
QY 242 QY-----ETSNPTVQLKIPOLRVSSVSKSQPDGSGGLDWMYQVSKTSSYLEGSAI 291
Db 928 TYSKIQDSKTDLIEDITPTFKKLITEMWENFMOLDLIIKIGIAKHCDLSSKYTVIITHXAC 987
QY 292 OKLKNILPKONKIESSGCVTHSSVDSYFLHGLSLP-----CLNSKNGIYDGTISEMT 343
Db 988 EKNLNVANSQNLVYTAETOIFDPDGT-----GNNSPILNIINDPTCONDENRCTEGTSNDN 1042
QY 344 EDGLDRKDSKOPRKKRGRYROY--DHEIMEE 372
Db 1043 EKCTIRSDCNSDKMEVFKLDGYPSDYDPPEE 1073

```

RESULT 10  
T50395  
actin-related protein [improved] - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
C:Accession: T50395  
R:Beck, A.; Borzjyn, K.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, January 1999  
A:Reference number: 225067  
A:Accession: T50395  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Reids: 1-433 <BEC>  
A:Cross-references: EMBL:AL116535; PIDN:CA86436.1; GSPDB:GN00067; SPDB:SPBP23A10.08  
A:Experimental source: strain 972h(-); clone p1 p23A10  
C:Genetics:  
A:Gene: SPDB:SPBP23A10.08  
A:Map position: 2  
A:Superfamily: actin

[illegible]

```

QY      301  ---QNKIECSGPVTHSSVSUYFLHGDLSPLCLNSKGTVD----GTSENTEDGIDR 349
          :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      374  YPGSRFLKIHASGVHVERSYASWLGSSILSSL-----GTFHQLMISRQVEEHSR 424

```

RESULT 11  
 A71928  
 cag island protein - *Helicobacter pylori* (strain J99)  
 C:Species: *Helicobacter pylori*  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: A71928  
 R:Alt, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 Ives, C.; Gibson, R.; Metberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A:Reference number: A71800; MUID:99120557; PMID:9933682  
 A:Accession: A71928  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1819 <ARN>  
 A:Cross-references: GB:AE001481; GB:AE001439; NID:g4155005; PIDN:ADD06047.1; PID:g4155000  
 A:Experimental source: strain J99  
 C:Genetics:  
 ;Gene: orf13/14

```

Query Match 15.6%; Score 125.5; DB 2; Length 1819;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 102; Conservative 70; Mismatches 180; Indels 165; Gaps 18;

Oy 14 SKSGKTOENRNGSIGPISVCKSIOMNO-----AENSLOEEOGPLDLYNRQEOQNTQQ 67
Db 72 SGGNETSETSSNGSIADLKLFFKKARKLYDCKRPFTQOKSLDETO---KLNEDEQENNEHQ 128
Oy 68 GGDVLIDSTKKTSTKSEESSICDPSSSENSVAGRHLNRREDYVERSAFAFGLLSKALKDI 127
Db 129 EETOTDLIDETSEKTKQODSPODLSNEATEA---NHFDLKLKSTESSNNHLDN----- 180
Oy 128 QSGALDINKAGILYGIPOKTLHLLEALPAGKPASFKNK-----TRDFHDSYSYKD 178
Db 181 -----PTESSDNHLDPETKTKQETHTHDEDDKPEEITODSDNQDEIKG 224
Oy 179 SKF-----TCVLOKVALMARA----- 195
Db 225 SKKKYIIIGIVAVLVIIILFSSRIFHYFVPLEDKSRFSKDRNLVYNDEIQIOREYNRL 284
Oy 196 QAERTESKSL-----NLLETSEI--KFTASTY-----LHOL 225
Db 285 LKENENEGNNIDIKLFFNDPNRTLYNYLMAIEDNGPLRAYECSNGNGNEECIKLI 344
Oy 226 TLQKMTQFEKNEESLYETSNPTVOLKIPOLRAVSSYSKQPDGSGLLDMVQVSKTSYV 285
Db 345 KDKKLQDMQKKTLEAVNDICKN-----AKTEBERIKCLDICKENLKSL 369
Oy 286 LEGS-----ALQXKNLIPRONKIECGPVTHSSVDSYF-----LHGDLPL--CLNS--- 331
Db 390 LNOQKQVAVLDCLNKMATDERKECKLKLINDPEIREKFRKLELOKLEOYKOCIKNAKT 449
Oy 332 ---KNGTVDGTSNTEDGDLR-----KDSKOPKKRGRROYDHEIMEEALA--NYMSG 380
Db 450 EAEKNECKLKLKSEALEIRLKLQALDCLKNKATDERKECKLKNIPDLOKELLDMSVKAY 509
Oy 381 KMSYSKAGIYGVPHSTLEYKVKERSGTLTPPKKYL 417
Db 510 KDCVSRAR-----NEKEQOBECKULTPEAKKL 536

RESULT 12
T41496
conserved hypothetical protein SPCC622.16c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
i:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
j:Accession: T41496

```

R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z21998

A:Accession: T14156

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-948 <SE>

A:Cross-references: EMBL:AL031127; PIDN:CAA21872.1; GSPDB:GN00068; SPDB:SPCC622.16c

A:Experimental source: strain 972h-, cosmid c622

C:Genetics:

A:Gene: SPDB:SPCC622.16c

A:Map position: 3

Query Match 5.6%; Score 125; DB 2; Length 948;

Best Local Similarity 19.6%; Pred. No.3.9; Mismatches 198; Indels 110; Gaps 19;

Matches 93; Conservative 74; Mismatches 198; Indels 110; Gaps 19;

Qy 7 QFAIEYISKSGK--TOENRNGSIGPSIVCKSIOMNQ-----AENSLQEOE--GFLD 54  
 Db 514 EICVDFVQKFGAMITVYHHRSAKHSCNCFSLQTKLIDSGPKPANNVYQHQSNIQVVI 573  
 Qy 55 LT---VNRMOBONTQGDGVLDTL-----STKKT-----SISESSICDPSE 94  
 Db 574 STNNHIIKKCOESQITGKNKNSFQIVKRIKSTKAPSWRSIIKAFKRENTRCNFISS 633  
 Qy 95 NSVAGRLHRN--REDYERSAEPADGLSKALKDIOGALDINKAGILYGIPOKTLHL 152  
 Db 634 -----LHATTFREDIVRPKIKSFVLQILFQALPFAIMWTPSPFNHNSFETALSK 687  
 Qy 153 EALPAGKPAFSPKNTKTRDPHDSY---SYKDSKETCAVLQKALMARQAERTKSKLTLE 209  
 Db 688 ETFNNGGAGNENTDTLTFTWGDGQFRPSDIC-----YDNFNLE 729  
 Qy 210 TSEIKFPFPASTYVLTQKAVTQFKENKESIQYETSNPVQLKIPQLRVSSVSKSQPDG 269  
 Db 730 TAN--SDAASIHLELQPL-NAVNERVIDISQDTMPSTAL--DPRVTRVDSLPDF 782  
 Qy 270 SGLL-----DVMYGVSTSSVLESGSALQKLNIPKONKIECGPVTSHSSVDSYFLHGD 323  
 Db 783 SNLILSPSSNDSPFLDLDLSPSSNLLKQIOKQVLP-ONSLEFSVGEKKAAYSLTLT 841  
 Qy 324 LSPILCNSKNGTVDGTSENTEDGLDRKDSKOPRRKRGYROYDHEIMEAIAWWSG-- 380  
 Db 842 FSVKRLSMENEPDPTTKVPLKYNIIQHEMKAYRRKN-----DLEIDQHPASSSGISN 895  
 Qy 381 -----KMSVSKAGQIYVPHSTLEYKXKERSGTLKTPPKKLRIPDTGLYMTD 429  
 Db 896 GRNNKREVNLTKAENV-GI-----KKGRIMKNENNIVDFED 930

#### RESULT 13

T14156 kinesin-related protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000

C:Accession: T14156

R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.

Cell 91, 357-366, 1997

A:Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromoso

A:Reference number: Z17893; PMID:98028574; PMID:9363944

A:Accession: T14156

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2954 <MO>

A:Cross-references: EMBL:AF027728; NID:92586070; PID:92586071; PIDN:AAC60300.1

C:Genetics:

A:Gene: XCEMP-E

C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 5.5%; Score 124.5; DB 2; Length 2954;

Best Local Similarity 18.5%; Pred. No.19; Mismatches 176; Indels 119; Gaps 19;

Qy 10 IEYISK-SGKTQENRNGSIGPSIVCKSIOMNQAENSLOEBOEGPLDTVNRMOBONTQOG 68  
 Db 533 LQYLPKDSGDMABERKXSFKEKETSLOOQLOSKEEKELVQS-FELKIALEEQ----- 586  
 Qy 69 DGVLDLSTKTKTSIKSESSSICDPSSNSVAGRLHRNEDYERS--AEFADGLLSKALKD 126  
 Db 587 -----LSVAKXNLEWNTNSREHINAIEVQTVKEVVRKMSVLDGSGYNASND 636  
 Qy 127 IQSGALDINKAGILYIG-----IPQTLHL--LALPAGKPAFKNT--RDPHDSY 174  
 Db 637 LQDSVVGKRLSSSHDCIEHRKMLEQIYDLLEEFLENLKK--KSENKQSSSEQDFWESI 695  
 Qy 175 SYKDS--KETCAVLOKVALW-----ARQAERTKSKNLLETSS 211  
 Db 696 QLCALIAEKAMLEELALMRDNPNTIILENETIKREIADLERLXKNQETNEFIIE-K 754  
 Qy 212 EIKFPFPASTYVLTQKAVTQFKENKESIQYETSNPVQLKIPQLRVSSVSK----- 264  
 Db 755 ETQKEHAQLIHEIGSLKLVENAMENYONLEEDLETKTLKQEIQLAELRRADNLQ 814  
 Qy 265 -----SQPQSGLDVMYGVSTSSVLESGSALQKLNIPKONKIECGPVTSHSS 314  
 Db 815 KVENPDLVSMGDSSEKLCEBIFQLKQSLSDAE-----VTRDAQKES----- 858  
 Qy 315 VDSYFLHGDLSPLCLNSKNGTVDGTSENTEDGLDRKD-----SKOPRRKRGYROYDHE 368  
 Db 859 -----FLRENELEKEMEDTSNWYQKKAASFQKQLETSNKKWMEAD 905  
 Qy 369 IMEBAIAWWSGKMSVSKAQIYG--VPHS-----TLEYKERSGTLTKPPKKK 416  
 Db 906 LQKE-----LQSAFNEIVNLGGLAGKVPRLDLSVLELEKVSSEFSKQLEKALEEK 956

#### RESULT 14

JC5837 364K Golgi complex-associated protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 20-Jun-2000

C:Accession: J05837

R:Toki, C.; Fujiwara, T.; Schda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.

Cell Struct. Funct. 22, 565-577, 1997

A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec

A:Reference number: JC5837; PMID:98093490; PMID:9431462

A:Accession: J05837

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3187 <OK>

A:Cross-references: DBJ:D25543; NID:9516825; PIDN:BA05026.1; PID:9516826

C:Comment: This protein plays a role in the formation and maintenance of the characteris

C:Superfamily: giantin

P:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predict

F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 5.5%; Score 124.5; DB 2; Length 3187;

Best Local Similarity 21.3%; Pred. No.22; Mismatches 129; Indels 121; Gaps 20;

Matches 85; Conservative 64; Mismatches 129; Indels 121; Gaps 20;

Qy 34 KSIMQNAENSLOEBOEGPLDTVNRMOBON-----TOQGDVLDLSTKTSIKSEBS 86  
 Db 2197 KEIWESKAQETELQHOQK-----AYDKLQENKELMSQLEBAGOLYHDSKNELTGLSEBLK 2251  
 Qy 87 SICPPSE--NSVVG-RLHNREDYERSAEPADGLSKALKDIOGALDINKAGILYGI 143  
 Db 2252 SLKQOSTDLKNSLKECKEHEHN-----LEGITIKQOBADQN----- 2287  
 Qy 144 POKTLHLLEALPAGKPAFSPKNTKTRDPHDSYSYKDSKETCAVLQKVALMARQAERTKS 203  
 Db 2288 -----CKNCEQLLELDLPAASELTTR-LHDEINVE-----QXI----- 2330  
 Qy 204 KLNILETSEIKFPFPASTYVLTQKAVTQFKENKESIQYETSNPVQLKIPOL 257  
 Db 2321 -ISLISKEEALQVAILHQQHSEIKELDENLISQ--EEENLTLEENKRAVEKTNOL 2377

QY 258 R-----VSVSKSQPDGSGLDVWYQVSK--TSSVLEGSALQK 295  
DB 2378 TEALETIKKESLEQKAQUDSFVKSMSSLODDRDRIYSDVRQLEERHLVILEKDEL--IQ 2435  
QY 236 NILPKONKI--ECGSPVTHSSVDSYFLHGDLSPLCLNSKNGTVDGT---SENTEDGLDR 349  
DB 2436 DAAANNKXKEIRG-----LRGHMD--LNSNAKLDAELIQYRRDLNEVIT 2482  
QY 350 KDSKQPRKRGYROYDHEIMEAIAWMSGKMSVSKAQ 388  
DB 2483 KDSQQRLLAQLOQ--NKLRENCVYK--LEGRLKGEAE 2518

## RESULT 15

T16270  
hypothetical protein F35D11.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16270

R:Fulton, B.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid F35D11.

A:Reference number: Z18487

A:Accession: T16270

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Cross-references: EMBL:U29381; NID:g668214; PID:g668224; PIDN:AAA68757.1; CESP:F35D11.

A:Experimental source: strain Bristol NZ

C:Gene: CESP:F35D11.11

A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1250/1

Query Match 5.5%; Score 123.5; DB 2; Length 1827;  
Best Local Similarity 19.6%; Pred. No. 12;  
Matches 92; Conservative 71; Mismatches 206; Indels 101; Gaps 17;

QY 18 KTOENRNGSIGPSIVCKSIOMNOAENSLOE--EOEGPLDTVYVNM-----OEONTQOG 68  
DB 1028 QNSELKNKGEG-----LSEKNNEBKRIODLADQLREANKVYVNMNMKNVNLSEKKNELD 1082  
QY 69 DGVLDLSTKKTISKSEESSICDPSSSENSVAGRLHRNREDEVSAEFADGLSKALKDIO 128  
DB 1083 QNVVTLDTNK--VRQLEIQMDKAAKNEVSGDLTRKE-----HDAQSMLKQANE-Q 1131  
QY 129 SCALDINKAGILYGIPOKTLHLHLEALPRG---KPAEFNKTRDPFHDSYSYKDSKETCAV 185  
DB 1132 FRLTDLKVRKALODENORLVNDLATVKAPEVKRETSSKAIISDILDKYRSAEERANKGE 1191  
QY 186 LQKVAL-----WARAOAERTE---KSKNLLETSEIKFPTASTYVHQ-----TLQKVV 231  
DB 1192 LONQRLRSPLATVTLTKERQELKAKDSQDRLRDSQGRFEVOSKLANLOKSAVESLQNP 1251  
QY 232 TOFKKKNESIQY-----ETSNPTVOLKIPLQLEVVSSVSKSQPDGSGLDVWYQV 279  
DB 1252 SSNSRQNRSIYVDIPRAASSIGLNNESDEVPLRSSPSVAFADSSQNMQRAVDSMDVSSV 1311  
QY 280 SKTSSVLE-----GSAIOLKNILPKONKIECGPVTTHSSVDSYFLHGDLS 326  
DB 1312 GVTLEFLKRIEQLADNADLSDALEKAKDELORNEKLRADROMVIERVERQLVH----- 1366  
QY 327 LCLNSKNGTVDG--TSEN-----TEGDGLDKSKQPKRGRYROYDHEIMEAIAWMSG 380  
DB 1367 --ITEERNTIENRMTSQRQMYITNESSRSRHEHTISMKARISTLEHLREKESKLAHLR 1424  
QY 381 KMSVSKAQGIYGVPHSTLEYKVKERSGTLKTPPKKRLPLDTGLVYNMDS 430  
DB 1425 K-----EIEVLHGQLHDALESKEA-----TGIWGVQDS 1453

Search completed: October 28, 2003, 12:03:16  
Job time : 31.7879 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 16.0727 Seconds  
(without alignments)  
1293.234 Million cell updates/sec

Title: US-10-016-768A-8

Perfect score: 2250  
Sequence: 1 MKKMIROPALEYISKSQKTQ.....GLYNTDGTGSCKNKSKPV 442

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131.5	5.8	997	B1R1_SCHPO	O14064 schizosach
2	131.5	5.8	2492	ATRX_HUMAN	P46100 homo sapien
3	131	5.8	1296	BXG_CLOBO	O60393 clostridium
4	123	5.5	3210	CENF_HUMAN	P49454 homo sapien
5	122	5.4	1164	KEL1_YEAST	P38853 saccharomyc
6	122	5.4	1435	LTEL1_YEAST	P07866 saccharomyc
7	122	5.4	1938	MYHD_HUMAN	O90433 homo sapien
8	120	5.3	1170	SMC2_YEAST	P38989 saccharomyc
9	118	5.2	635	HS68_DROME	O91125 drosophila
10	118	5.2	1147	RFC1_HUMAN	P35251 homo sapien
11	117	5.2	1146	KHMA_DICDI	P42527 dictyosteli
12	117	5.2	1802	HKR1_YEAST	P41809 saccharomyc
13	117	5.2	1940	MYH3_HUMAN	P11055 homo sapien
14	117	5.2	2188	POLG_EC23C	O99148 e genome po
15	116	5.2	1258	NEK1_HUMAN	O96066 homo sapien
16	115.5	5.1	2156	RPI_HUMAN	P56715 homo sapien
17	115.5	5.1	2230	GOS4_HUMAN	O13439 homo sapien
18	115	5.1	1535	LMH1_CAEEL	Q18823 caenorhabdi
19	115	5.1	1940	MYH3_RAT	P12847 rattus norv
20	114.5	5.1	651	SEC9_YEAST	P40357 saccharomyc
21	114.5	5.1	2476	ATRX_MOUSE	O61687 mus musculu
22	114.5	5.1	2845	APC_MOUSE	O61315 mus musculu
23	113.5	5.0	1630	MSPI_PLARF	P04932 plasmodium
24	113.5	5.0	1639	MSPI_PLARF	P04933 plasmodium
25	113	5.0	1531	NFT5_HUMAN	O94916 homo sapien
26	113	5.0	1690	CJ90_DROME	O99416 drosophila
27	113	5.0	1875	MLP1_YEAST	O02455 saccharomyc
28	113	5.0	2017	MYSN_DROME	O99323 drosophila
29	112.5	5.0	944	NUP1_YEAST	P33380 saccharomyc
30	112.5	5.0	2004	MOZ_HUMAN	O92794 homo sapien
31	112	5.0	609	YSWI_YEAST	P38280 saccharomyc
32	111.5	5.0	1790	USO1_YEAST	P23386 saccharomyc
33	111.5	5.0	1939	MYH6_MESAU	P13539 mesocricetu

34	111.5	5.0	1972	MYH6_MOUSE	O08638 mus musculu
35	111.5	5.0	2291	SPCB_DROME	O00963 drosophila
36	111	4.9	633	BZL1_YEAST	P38822 saccharomyc
37	111	4.9	3924	ANK2_HUMAN	O01484 homo sapien
38	110.5	4.9	432	HMAT_BACSU	O07621 bacillus su
39	110.5	4.9	1938	MYH6_MOUSE	O02566 mus musculu
40	110.5	4.9	2116	MYH2_DICDI	O08799 dictyosteli
41	110	4.9	1233	SMC1_SCHPO	O94383 schizosach
42	110	4.9	1324	SMC4_SCHPO	P41004 schizosach
43	110	4.9	1939	MYH6_HUMAN	P13533 homo sapien
44	110	4.9	2179	POLG_EC23C	O73556 e genome po
45	109.5	4.9	688	LIP_STAEP	O02510 staphylococ

## ALIGNMENTS

RESULT 1  
B1R1\_SCHPO STANDARD; PRT; 997 AA.  
AC O14064:  
15-VUL-1998 (Rel. 36, Created)  
DT 15-VUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Birt protein (Chromosome segregation protein cut17).  
GN Birt OR CUT17 OR PBH1 OR SPC962.02C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N. A., FUNCTION AND SUBCELLULAR LOCATION.  
RX MEDLINE=21439264; PubMed=11554922;  
RA Morishita J., Matsusaka T., Goshima G., Nakamura T., Tatebe H.,  
RA Yanagida M.,  
RT "Birt/Cut17 moving from chromosome to spindle upon the loss of  
RT cohesion is required for condensation, spindle elongation and  
RT repair".  
RL Genes Cells 6:743-763(2001).  
[2]  
RP SEQUENCE FROM N. A.  
RP STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros K., Peac N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jags K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Roben J., Grymopoulos B.,  
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
RA Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Motier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hutz S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Uesery D., Bartell B.G., Nurse P.,  
RT "The genome sequence of Schizosaccharomyces pombe".  
RL Nature 415:871-880(2002).  
[3]  
RP CHARACTERIZATION.  
RX MEDLINE=99398681; PubMed=10468581;

```

RA Uren A.G., Beilharz T., O'Connell M.J., Bugg S.J., van Driel R.,
RA Vaux D.L., Lithgow T.;
RT "Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell
RT division.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10170-10175(1999).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=21850422; PubMed=11861551.
RA Rajagopalan S., Balasubramanian M.K.;
RT "Schizosaccharomyces pombe Bir1, a nuclear protein that localizes to
RT kinetochores and the spindle midzone, is essential for chromosome
RT condensation and spindle elongation during mitosis.";
RL Genetics 160:445-456(2002).
RN [5]
RP FUNCTION.
RX MEDLINE=20035862; PubMed=10571085;
RA Rajagopalan S., Balasubramanian M.K.;
RT "S. pombe Pbh1p, an inhibitor of apoptosis domain containing protein
RT is essential for chromosome segregation.";
RL FEBS Lett. 460:187-190(1999).
CC -!- FUNCTION: Seems to act in the pleiotropic control of cell
CC division. Has a role in chromosome segregation by recruiting
CC condensin and ark1 kinase to appropriate sites as the cell
CC progresses through mitosis.
CC -!- SUBCELLULAR LOCATION: Nuclear. Interacts with the outer
CC centromeric regions of the chromosomes during interphase. After
CC chromatin separation moves to the middle of the spindle.
CC -!- SIMILARITY: Contains 2 BIR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011034; BAA03415.1; -
DR EMBL; AL011323; CAA00434.1; -
DR PIR; T43523; T43523.
DR HSSP; Q13490; 10BH.
DR GenDB; SPombe; SPCC962.02c; -.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 2.
DR SMART; SM00238; BIR; 2.
DR PROSITE; PS01282; BIR_REPEAT_1; FALSE_NEG.
DR PROSITE; PS0143; BIR_REPEAT_2; 2.
KW Cell division; Mitosis; Nuclear protein; Repeat.
FT REPEAT 25 99 BIR 1.
FT REPEAT 120 194 BIR 2.
FT DOMAIN 80 83 POLY-ASP.
FT DOMAIN 312 319 POLY-ASP.
FT DOMAIN 487 490 POLY-SER.
SQ SEQUENCE 997 AA; 112579 MW; 952AGBAFA5C489F4 CRC64;
Query March 5.8%; Score 131.5; DB 1; Length 997;
Best Local Similarity 21.7%; Pred. No. 1.2; Indels 141; Gaps 24;
Matches 109; Conservative 77; Mismatches 175;

```

```

DB 570 TAIHVSKEFDLENKSMSEBSQSLQSLSEENDDKPLDILDLPLAIK-----RKDN 618
QY 230 WYTFQFKKMS-----LQYETSNPTVOLKTPQLRVSSV-----SKSPDSSGLDWWYQ- 278
DB 619 LVSGVLEKKGSTSTSKKPTSTIVDF-IEKPKTEISEVLPEKKRAICDESSQTVRVSIDR 677
QY 279 -VSKTSSVLGSLALOKLKNLIPKONKIECGPVT-----HSSVQ-----SY-- 318
DB 678 GVTYTRVSSPVSDEKENV-----NHEANSGHITWNVHSSLDPPQIVQNELESGLYK 733
QY 319 -----FLHGLSPPLCNLSKNG-TYDGTSENTEDGLDRKDSKQPKKGRYR 363
DB 734 DLPPRVNGSEKVTFGEDDINSPLQSKNNQTVAVTETSDKLQEKXA----- 782
QY 364 QYDHEI-----MEALMVMSGKMSVSKAGITGVPHSTLEYKVKESGLTKTPPK-KL 417
DB 783 --NHELENIKEIEKLETV--DKVYSLDAPDDEIKNSRTSVNGTRSVSKNTPKTKV 838
QY 418 RLDPDTGLYNNMTDSGTGCKNKS 439
DB 839 DKIDNVSKDVEVTPSGSCETSS 860

RESULT 2
ATTRX HUMAN
ID ATTRX HUMAN STANDARD; PRT: 2492 AA.
AC P46100; F51068; Q15886; Q9H021; Q9NTS3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcriptional regulator ATTRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP) (Znf-HX).
GN ATTRX OR RAD54L OR XH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT SER-1660, AND
RP VARIANTS ATR-X.
RX MEDLINE=97123494; PubMed=8968741;
RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,
RA Gibbons R.J.;
RT "ATTRX encodes a novel member of the SNF2 family of proteins: mutations
RT point to a common mechanism underlying the ATR-X syndrome.";
RL Hum. Mol. Genet. 5:1899-1907(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RX MEDLINE=97386582; PubMed=924431;
RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chiaroni P.,
RA Colliex L., Schwartz C., Fontes M.;
RT "Determination of the genomic structure of the XNP/ATTRX gene encoding
RT a potential zinc finger helicase.";
RL Genomics 43:149-155(1997).
RN [3]
RP SEQUENCE OF 860-2492 FROM N.A.
RX MEDLINE=95179111; PubMed=7874112;
RA Stayton C.L., Dobovic B., Guisano M., Geetz J., Broccoli V.,
RA Giovannazzi S., Bosso Lasco M., Monaco L., Rastan S., Boncinelli E.,
RA Bianchi M.E., Consalez G.G.;
RT "Cloning and characterization of a new human Xq13 gene, encoding a
RT putative helicase";
RL Hum. Mol. Genet. 3:1957-1964(1994).
RN [4]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=94214473; PubMed=8162050;
RA Geetz J., Pollard H., Consalez G., Villard L., Stayton C.L.,
RA Millaesau P., Khrestchatsky M., Fontes M.;
RT "Cloning and expression of the murine homologue of a putative human
RT X-linked nuclear protein gene closely linked to Pkci in Xq13.3.";
RL Hum. Mol. Genet. 3:39-44(1994).
RN [5]
RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.

```

RX MEDLINE=95211835; PubMed=7697714;  
 RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;  
 RT "Mutations in a putative global transcriptional regulator cause X-  
 RL linked mental retardation with alpha-thalassemia (ATR-X syndrome).";  
 RL Cell 80:837-845(1995).  
 RN [6]  
 RN SEQUENCE OF 1375-2492 FROM N.A.  
 RA Pearce A., Chapman J.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN EZH2 BINDING.  
 RA MEDLINE=98167853; PubMed=9499421;  
 RA Cardoso C., Timsit S., Villard L., Khrestchatskiy M., Fontes M.,  
 RA Colliaux L.;  
 RT "Specific interaction between the XNP/ATR-X gene product and the SET  
 RT domain of the human EZH2 protein.";  
 RL Hum. Mol. Genet. 7:679-684(1998).  
 RN [8]  
 RN SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC  
 RP HETEROCHROMATIN.  
 RA MEDLINE=20040663; PubMed=10570185;  
 RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,  
 RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,  
 RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;  
 RT "Localization of a putative transcriptional regulator (ATRX) at  
 RT pericentromeric heterochromatin and the short arms of acrocentric  
 RT chromosomes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13986(1999).  
 RN [9]  
 RN DISEASE.  
 RA MEDLINE=20213147; PubMed=10751095;  
 RA Villard L., Fontes M., Ades L.C., Geetz J.;  
 RT "Identification of a mutation in the XNP/ATR-X gene in a family  
 RT reported as Smith-Fineman-Myers syndrome.";  
 RL Am. J. Med. Genet. 91:83-85(2000).  
 RN [10]  
 RN VARIANT ATR-X SER-1713.  
 RA MEDLINE=97196774; PubMed=9043863;  
 RA Villard L., Lacombe D., Fontes M.;  
 RT "A point mutation in the XNP gene, associated with an ATR-X phenotype  
 RT without alpha-thalassemia.";  
 RL Eur. J. Hum. Genet. 4:316-320(1996).  
 RN [11]  
 RN VARIANT JM GLN-2131.  
 RA MEDLINE=96624392; PubMed=8630485;  
 RA Villard L., Geetz J., Mattei J.-F., Fontes M., Saugier-Verber P.,  
 RA Munnich A., Lyonnet S.;  
 RT "XNP mutation in a large family with Jubb-Marsidi syndrome.";  
 RL Nat. Genet. 12:359-360(1996).  
 RN [12]  
 RN VARIANTS ATR-X.  
 RA MEDLINE=97467722; PubMed=9326931;  
 RA Gibbons R.J., Bachoo S., Picketts D.J., Afimos S., Asehnauer B.,  
 RA Bejoffon U.J., Berry S.A., Dahl N., Fryer A., Keppeler K., Kurosawa K.,  
 RA Levin M.L., Masuno M., Neri G., Pierpont M.E., Slaney S.F.,  
 RA Higgs D.R.;  
 RT "Mutations in transcriptional regulator ATRX establish the functional  
 RT significance of a PHD-like domain.";  
 RL Nat. Genet. 17:146-148(1997).  
 RN [13]  
 RN VARIANT ATR-X LEU-246.  
 RA MEDLINE=20123062; PubMed=10660327;  
 RA Fichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S.,  
 RA Greco D., Cardoso C., Fontes M., Ragusa A.;  
 RT "New mutations in XNP/ATR-X gene: a further contribution to  
 RT genotype/phenotype relationship in ATR/X syndrome.";  
 RL Hum. Mutat. 12:214-214(1998).  
 RN [14]  
 RN VARIANT SHS LYS-1742.  
 RA MEDLINE=99347960; PubMed=10417298;  
 RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,  
 RA Prieto F., Fontes M., Martinez F.;  
 RT "Mutation of the XNP/ATR-X gene in a family with severe mental

RT retardation, spastic paraplegia and skewed pattern of X inactivation:  
 RT demonstration that the mutation is involved in the inactivation  
 RT bias.";  
 RL Am. J. Hum. Genet. 65:558-562(1999).  
 RN [15]  
 RN VARIANT CWS THR-2050.  
 RA MEDLINE=99326061; PubMed=10398237;  
 RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,  
 RA Curtis M.;  
 RT "Carpenter-Maziri syndrome results from a mutation in XNP.";  
 RL Am. J. Med. Genet. 85:249-251(1999).  
 RN [16]  
 RN VARIANTS ATR-X GLU-175; 178-VAL--LYS-198 DEL; SER-190; PRO-219;  
 RP LEU-246 AND CYS-249.  
 RA MEDLINE=99719535; PubMed=10204841;  
 RA Villard L., Bonino M.-C., Abidi F., Ragusa A., Belongue J.,  
 RA Lossi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M.,  
 RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;  
 RT "Evaluation of a mutation screening strategy for sporadic cases of  
 RT ATR-X syndrome.";  
 RL J. Med. Genet. 36:183-186(1999).  
 RN [17]  
 RN VARIANTS ATR-X SER-179; LEU-190; ILE-194; CYS-246; PHE-1552; SER-1645  
 RP AND CYS-1847.  
 RA MEDLINE=20451413; PubMed=10995512;  
 RA Wada T., Kubota T., Fukushima Y., Saitoh S.;  
 RT "Molecular genetic study of Japanese patients with X-linked alpha-  
 RT thalassemia/mental retardation syndrome (ATR-X).";  
 RL Am. J. Med. Genet. 94:242-248(2000).  
 CC -I- FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES  
 CC GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN BRAIN  
 CC DEVELOPMENT AND FACIAL MORPHOGENESIS.  
 CC -I- SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND  
 CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSERINE-DEPENDENT MANNER (BY  
 CC similarity).  
 CC -I- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC  
 CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY  
 CC INTERACTING WITH HP1.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=5;  
 CC Name=4;  
 CC IsoId=P46100-1; Sequence=Displayed;  
 CC Name=1;  
 CC IsoId=P46100-2; Sequence=VSP\_000575;  
 CC Name=2;  
 CC IsoId=P46100-3; Sequence=VSP\_000574;  
 CC Name=3;  
 CC IsoId=P46100-4; Sequence=VSP\_000576;  
 CC Name=5;  
 CC IsoId=P46100-5; Sequence=VSP\_000574, VSP\_000576;  
 CC -I- TISSUE SPECIFICITY: Ubiquitous.  
 CC -I- DISEASE: Defects in ATRX are the cause of X-linked alpha-  
 CC thalassemia/mental retardation syndrome (ATR-X) [MIM:301040]. ATR-  
 CC X is an X-linked disorder comprising severe psychomotor  
 CC retardation, facial dysmorphism, urogenital abnormalities, and  
 CC alpha-thalassemia. An essential phenotypic trait are hemoglobin H  
 CC erythrocyte inclusions.  
 CC -I- DISEASE: Defects in ATRX are the cause of Sutherland-Haan X-linked  
 CC mental retardation syndrome (SHS) [MIM:309470]. It is  
 CC characterized by severe mental retardation and cryptorchidism.  
 CC parasplesia, microcephaly, short stature and cryptorchidism.  
 CC -I- DISEASE: Defects in ATRX are a cause of Smith-Fineman-Myers  
 CC syndrome (SFM) [MIM:309580]. Clinical features include severe  
 CC mental retardation, microcephaly, growth failure, facial anomalies  
 CC and bilateral cryptorchidism. Due to the clinical overlap with  
 CC ATR-X syndrome, some patients originally diagnosed as having SFM,  
 CC might be affected by a variant of ATR-X syndrome which lack  
 CC hemoglobin h inclusions.  
 CC -I- DISEASE: Defects in ATRX are the cause of Carpenter-Maziri  
 CC syndrome (CWS), an X-linked recessive condition characterized by  
 CC moderate mental retardation, short stature, brachydactyly with  
 CC excessive skin creases, and widening of the knuckles.  
 CC -I- DISEASE: Defects in ATRX are the cause of Jubb-Marsidi syndrome

(JMI) [MIM:309590]. JM is a rare X-linked recessive disease characterized by severe mental retardation, growth failure, sensorineural deafness, microgenitalism and early death.

-1- SIMILARITY: BELONGS TO THE SNF2/RAD50 HELICASE FAMILY.

-1- SIMILARITY: Contains 1 PHD-type zinc finger.

Query Match 5.8%; Score 131.5; DB 1; Length 2492;  
Best Local Similarity 23.0%; Pred. No. 4.1; Indels 99; Gaps 21;  
Matches 100; Conservative 63; Mismatches 173;

14 SKSGTQNRNG-SIGPSIVCKSIOMNOAENSIOEEOGPDLDITVNRMOEQTOOGDQVL 72  
DB STSGSDPTTKGKSKSSIISSKKKQTOSES---NYNSELEKEKSKSGAAR----- 835

73 DLSTK--TSIXSESSICDPSSSENSVAGRLHNRNEDYVERSAEPADGLSKALDKIOSG 130  
DB --TTKRIPTNTPDPSSEDEKSKKGMNQGKHLKTKSQEGSSDPAERKQERETPSAEG 893

131 ALDINKAGILVIGIPKXTLLHL--EALPAKPKASFKNKRDRPHDSYSYDSEKTCVLOKV 189  
DB --KDTIMELRDRLPKKQAS--ASTDGVDKLSGKEQSFSLSEVRKV 939

190 ALMARAQAERTKSKLNTLETSEIKPTASTYLIHQLTKMVTQFKENESIQVETSNPT 249  
DB --AETKSKK-----HLKTKCKV--ODGSLDAEKLLKDDQS--DETSED 981

250 VOLKIPQIRVSSVSKSPDGGSLDVMYQVSKTSSVLEGSALQKLKNTLPKONKIECSGP 309  
DB --QSKKGTEEKKPS-----DFKKKVIEMEQVSSSGTEK--LPEREEI--CHFP 1029

310 VTHSSVDVYFLHGLDPLCLNSKNGTVDGTSNTEGDLDRKDSKPKKRGRIYRDHEI 369  
DB --KNGTGDG-----EKSKKIRDKTSKKDELSDY 1063

1030 KGIKOI-----KNGTGDG-----EKSKKIRDKTSKKDELSDY 1063

370 MEEAIAWMSGKMSYSK--AAGIYVPHSTLEYKVERSGTLKTPPKK--LRLPDGLY 425  
DB 1064 AEKSTGKSDSDSSDKSKXGAYG-----REKKCKLKGSKRRKQDCSSSDTKY 1115

426 NMTDSGTGSCKNSSK 440  
DB 1116 SMKEDG---CNSSDK 1127

RESULT 3  
EXG\_CLOBO STANDARD; PRT; 1296 AA.

AC Q60393;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)  
GN BONTG.  
OS Clostridium botulinum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1491;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX STRAIN=113 / 30;  
RX MEDLINE=94092745; PubMed=8268233;  
RA Campbell K., Collins M.D., East A.K.;  
RT "Nucleotide sequence of the gene coding for Clostridium botulinum  
RT (Clostridium argentinense) type G neurotoxin: genealogical comparison  
RT with other clostridial neurotoxins."  
RL Biochim. Biophys. Acta 1216:487-491 (1993).  
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
CC ENDOPEPTIDASE.

-1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
CC detected action on small molecule substrates.  
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
CC heavy chain (H). The light chain has the pharmacological activity,  
CC while the N- and C-terminal of the heavy chain mediate channel  
CC formation and toxin binding, respectively.  
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL; X74162; CA52275.1; -.  
DR HSSP; P10845; 3BTA.  
DR MEROPS; M27.002; -.  
DR InterPro; IPR000395; Bontoxilysin.  
DR InterPro; IPR006025; Zn\_MTPepidase.  
DR Pfam; PF01742; Peptidase\_M27; 1.  
DR PRINTS; PR00760; BONTOTOXILYSIN.  
DR PRODOM; PD001963; Bontoxilysin; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Hydrolase; Metalloprotease; Zinc.  
FT INT MET 0 441  
FT CHAIN 1 441  
FT CHAIN 442 1296  
FT METAL 229 229  
FT ACT SITE 230 230  
FT METAL 233 233  
FT DISULFID 435 449  
SQ SEQUENCE 1296 AA; 149013 MW; DCEB47E15F665C31 CRC64;

Query Match 5.8%; Score 131; DB 1; Length 1296;  
Best Local Similarity 20.2%; Pred. No. 1.8;  
Matches 78; Conservative 64; Mismatches 111; Indels 114; Gaps 15;

25 GSIGPSIVCKSIOMNOAENSIOEEOGPDLDITVNRMOEQTOOGDQVLDSTKTSIXSE 84  
DB 272 GGHDPVSISPSTDNINIKALQNPQD-----IANRLNIVSSAQSGI-DISLVKQIYK 325

85 ESSICDPSSSENSVAGRLHNRNEDYVERSAEPADGLSKALDKIOGALDINKAGILYGP 144  
DB 326 YDFVEDPVGKISV-----DKDKF-----DKLYKALMFGFTETNLAG-EYGI- 365

145 QKTLLHL--EALP-----AGKPAFPKTRDPF-----DSYSYKDSKE 181  
DB 366 -KTRYSFSEVLPRTKTEKLDNTITQNGFPNLSKNTKTEFGQKAVKAEYEEISL 424

182 TCAYLQVALMARAQAERTKSKLNTLETSEIKPTASTYLIHQLTKMVTQFKENESL 241  
DB 425 EHLVIYRIAMCKPVMYKNTGSEECIIVNNEDEFFIAN-----XDSESKDLAKAETI 476

242 QVETSNPTVQ-----LQIP-QIRVSSVSKSP 267  
DB 477 AYNQNNNTIENFESIDQLIDNDLSSIDLPLENTEFTNFDIDIVYIKQSLAKKIFV 536

268 DSGSLDVMYQVSKTSSVLEGSALQKLKNTLPKONK-----IECSGPHTHSSVDSY 318  
DB 537 DGDSDLFYLHAQTPPSNIENLQNLNSLDALRRNNKKYTFSTLVLEKANTVVGAS----- 592

319 FLHGDLSPLCLNSKNGTVDG-TSENTE 344  
DB 593 -----LFWNVKGVVIDFTSESTQ 611

RESULT 4  
 CENF HUMAN STANDARD: PRT; 3210 AA.  
 ID CENF HUMAN 013171; 013246;  
 AC P49454; 013171; 013246;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE CENP-F kinetochore protein (Centromere protein F) (Mitosis) (AH antigen).  
 GN CENP-F.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=95348175; PubMed=7542657;  
 RA Liao H., Winkler R.J., Mack G., Rattner J.B., Yen T.J.;  
 RT "CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";  
 RL J. Cell Biol. 130:507-518(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=95379848; PubMed=7651420;  
 RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B., Jones D., Yang-Feng T.L., Lee W.-H.;  
 RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein that is specifically involved in mitotic-phase progression.";  
 RL Mol. Cell. Biol. 15:5017-5029(1995).  
 [3]  
 RP SEQUENCE OF 2194-3210 FROM N.A.  
 RX MEDLINE=9536446; PubMed=7612011;  
 RA Li Q., Ke Y., Kapp J.A., Fertis N., Medsger T.A. Jr., Joshi H.C.;  
 RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal domain sufficient for nuclear localization.";  
 RL Biochem. Biophys. Res. Commun. 212:220-228(1995).  
 [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95370296; PubMed=7642639;  
 RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and HUBB1.";  
 RL J. Cell Biol. 143:49-63(1998).  
 [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98437347; PubMed=9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and HUBB1.";  
 RL J. Cell Biol. 143:49-63(1998).  
 [6]  
 RP FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH KINETOBLASTOMA PROTEIN (KB), CENP-E AND HUBB1.  
 CC - SUBUNIT: HOMO- OR HETERODIMER.  
 CC - SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS). REORGANIZATION TO THE KINETOCHORE/CENTROMERE (CORONAL SURFACE OF THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.  
 CC - DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.  
 CC - PTM: HYPERPHOSPHORYLATED DURING MITOSIS.  
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; U19769; AAA82889.1;  
 CC EMBL; U30872; AAA82935.1;  
 CC EMBL; U25725; AAA86889.1;  
 CC PIR; PC4035; PC4035.

DR Genew; HGNC:1857; CENPF.  
 DR GK; P49454; --  
 DR MIM; 600236; --  
 DR GO; GO:0005699; C.kinetochore; TAS.  
 DR GO; GO:0005634; C.nucleus; TAS.  
 DR GO; GO:0005819; C.spindle; TAS.  
 DR GO; GO:000067; P.DNA replication and chromosome cycle; TAS.  
 DR GO; GO:0007088; P.regulation of mitosis; TAS.  
 KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil; Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat; Polymorphism.  
 FT DOMAIN 14 197  
 FT DOMAIN 273 769  
 FT DOMAIN 823 1328  
 FT DOMAIN 1642 1746  
 FT DOMAIN 1862 2987  
 FT DOMAIN 2207 2568  
 FT REPEAT 2207 2386  
 FT REPEAT 2389 2568  
 FT DOMAIN 3015 3032  
 FT VARIANT 3202 3202  
 FT CONFLICT 16 16  
 FT CONFLICT 250 250  
 FT CONFLICT 272 272  
 FT CONFLICT 611 611  
 FT CONFLICT 1494 1589  
 FT CONFLICT 1611 1611  
 FT CONFLICT 1811 1811  
 FT CONFLICT 2242 2243  
 FT CONFLICT 2335 2335  
 FT CONFLICT 2492 2492  
 FT CONFLICT 2545 2561  
 SQ SEQUENCE 3210 AA; 367589 MW; 11D83324960E4334 CRC64;  
 Query Match 5.5%; Score 123; DB 1; Length 3210;  
 Best Local Similarity 20.0%; Pred. No. 19;  
 Matches 103; Conservative 74; Mismatches 20; Indels 138; Gaps 21;  
 7 GPATYISKSGKTOENR--NGSICPSIVCKSIOMNOENSIQEOBPLDITVRMOQ 63  
 2676 QDTLEVLQSSYNNLENELETKMDKMSFEVKNTAKETELQRMHMAQKTAELOEL 2735  
 64 NNGQSDGVLDLSTKTSIKSESSICDPSSNS-VAGLHNRREYVERSAFAAGL--- 119  
 2736 SGEKRLAGEQLLEIEIKSSDOQLKELTLENSLKSLDCMKHQVKEGKVRREILAY 2795  
 120 ---LSKALDKIQSGALDINKAGILYGIPOKTLHLLEALPAGKPASFKNKTDFHDSYS 176  
 2796 QLRLEAKKQALLDITNKO---YEVEIOT-----YREKL----- 2828  
 177 KDSKTCVAVLOKVALMARAQERTKSKLN--LLETSSIKFPTASTYVHLQTLQKMTQF 234  
 2829 -TSKEECSSQKLEI-----DLKSSKEELNNSLKATTOILELKKTKMDNL---KYVQL 2880  
 225 KEKNSL-----QYTSNPYQAKIPQARVSVSKSGDGGGLDVMQVSKT 282  
 2881 KKENRAQKKMLIKSKQLEBEKEILQKELSOLAQO-----EKQT 2924  
 283 SSYLE-----GSALQKLNILPKONKIE-----CGPVTHSSVD---SYFLHGLDPL 327  
 2925 GTVMQTKYDELTEYKEKLEKTEKTEADEYLDKCYCLLSHETLEBAKKELETQVAHL 2984  
 328 C-----LNSKNGTVDG-----TSENTEGDLDRKSKOPRKK-----RGARYO 364  
 2985 CSQSKQSDRSRSPPLGVPVPPSPPIPVTEKRLSGGKAGKGRORSSGIWENGCGPTPA 3044  
 365 YDHEIMEAIAIMWSGKGSVSKAGCI---YGVPH-----STLEKVKVERGCT 408  
 3045 TPESFSKSKKAVWSGHPADBTETETEPFGGLPVPVKKGADIPTKGTSPIILRTTMA 3104  
 409 LKTPPK---KKLRLEPDTGLYNNMTDGTGSCSKNSKP 441

DB 3105 TRTSPRLAAOKLALSPSL-----GKENVLAESSKP 3134

## RESULT 5

ID\_KEL1\_YEAST STANDARD; PRT; 1164 AA.  
 AC P38853;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DE Kelch repeats protein 1.  
 GN KEL1 OR YHR158C.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 CX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 Kučaba T., Hillier L., Jier M., Johnston L., Langston Y.,  
 Lacroix P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,  
 Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 Vaudin M.,  
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 VII.";  
 RL Science 265:2077-2082 (1994).  
 RP CHARACTERIZATION.  
 RX MEDLINE=99003296; PubMed=9786949;  
 RA Phillips U., Herskowitz I.,  
 RT "Identification of Kel1p, a kelch domain-containing protein involved  
 in cell fusion and morphology in Saccharomyces cerevisiae.";  
 RL J. Cell Biol. 143:375-389 (1998).  
 CC -!- FUNCTION: HAS A ROLE IN CELL MORPHOGENESIS AND CELL FUSION AND MAY  
 ANTAGONIZE THE PKC1 PATHWAY.  
 CC -!- SUBUNIT: INTERACTS WITH KEL2.  
 CC -!- SIMILARITY: Contains 5 kelch repeats.  
 CC -!- SIMILARITY: TO YEAST KEL2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, U10397; AAB68991.1; -.  
 DR PIR, S46769; S46769.  
 DR COMPUYEAST-2DPAGE; P38853; -.  
 DR SCD, S0001201; KEL1.  
 DR GO: GO:0005935; C:bud neck; IDA.  
 DR GO: GO:0005934; C:bud tip; IDA.  
 DR GO: GO:0005737; C:cycloplasm; IDA.  
 DR GO: GO:0005937; C:shmoo tip; IDA.  
 DR GO: GO:0000755; P:cyclogamy; IGI.  
 DR GO: GO:0008360; P:regulation of cell shape; IMP.  
 DR InterPro: IPR006552; Kelch\_rep.  
 DR Pfam: PF01344; Kelch\_4.  
 KM Kelch repeat; Repeat; Coiled coil.  
 FT REPEAT 139 186 KELCH 1.  
 FT REPEAT 253 307 KELCH 2.  
 FT REPEAT 308 357 KELCH 3.  
 FT REPEAT 359 409 KELCH 4.  
 FT REPEAT 411 460 KELCH 5.  
 FT DOMAIN 777 931 COILED COIL (POTENTIAL).  
 FT DOMAIN 974 1163 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 1164 AA; 131093 MW; 43D0F570F1D54D CRC64;

Query Match 5.4%; Score 122; DB 1; Length 1164;  
 Best Local Similarity 19.7%; Pred. No. 5.6;  
 Matches 85; Conservative 69; Mismatches 159; Indels 118; Gaps 17;

QY 7 QPAIEYISKSGKTOENNGSIGPSIVCKSIQMOAENSIOEBGPDLTIVNRQEQNTQ 66  
 DB 694 QFKIKHYNESELSQN-----NTEIDKLE-----PVDITKKSDTGAGHD 733  
 QY 67 QGDGVDLSTTK-----TSIKSESSICDPSSSNSVAGRLHRREDYVERSAEFAD 117  
 DB 734 SANHVIVASDEKXVSPMGDVPPTDKNEASV--PINDATV-----EVDRA----- 778  
 QY 118 GLSKALKDIOGALDINKAGILYGIPOKTLHLLELPAGKPAFPKXTRDFHDSYSYK 177  
 DB 779 -LFEKLASELOS-----LKELTHERKLEAG--AHTELETELMOQLSKQ 819  
 QY 178 DSKETCAV--LQKVALARAQERTKSKLNLLETSEIKFPYASTYVJHQLQK----- 230  
 DB 820 NSGTTEIDELDSVRL-----QSKCEILEADNHSLEDKVNLELELVNSKFLDIEN 869  
 QY 231 ---VTQK--EKNESLQYETSNPTVQLKIPQLRVSSVSKQPDGGLDWVQVSKTSSVL 286  
 DB 870 LNEVIOFQNEKIKSLLE---PNYKELLELQLEHENLSREN----- 908  
 QY 287 EGSALQKXKNTLPKONKIEGSPVTHS---SVDSYFLHGDLSPLCLNSKXGTVDTGSEN 342  
 DB 909 -----ERLKNEKSHNEDIINNVANYSQGLSLSHKERNANSFLESSSLISVDEN 963  
 QY 343 TEDGLDRKDSKOPRKE---GRYQYDHEIMEEAIAMWSGKNSVSKAOGIYGVPHSTLE 399  
 DB 964 GEKTVGEPPYGQSRHHRVINKLTNRDLRLERQELTIS-KEXLSSEYHALKMEHSSLIS 1022  
 QY 400 YKVERBSGTLK 410  
 DB 1023 QDVLVKNENIK 1033

## RESULT 6

ID\_LTEL\_YEAST STANDARD; PRT; 1435 AA.  
 AC P07866;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Low temperature essential protein.  
 GN LTEL OR MS12 OR YAL024C.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 CX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=95249563; PubMed=7985422;  
 RA Keng T., Clark M.W., Storms R.K., Fortin N., Zhong W.,  
 Roulette F.B.F., Barton A.B., Kaback D.B., Bussey H.,  
 RT "LTEL of Saccharomyces cerevisiae is a 1435 codon open reading frame  
 that has sequence similarities to guanine nucleotide releasing  
 factors.";  
 RL Yeast 10:953-958 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=95249563; PubMed=7731988;  
 RA Bussey H., Kaback D.B., Zhong W.,  
 RT "The nucleotide sequence of chromosome I from Saccharomyces  
 cerevisiae.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813 (1995).  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95028143; PubMed=7941731;





the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U05820; AAA17416.1; -  
 CC EMBL; D50617; BAA09270.1; -  
 DR PIR; A56157; A56157.  
 DR SGD; S0001927; SMC2.  
 DR GO; GO:0005676; C:condensin complex; IPI.  
 DR GO; GO:0004002; P:adenosinetriphosphatase activity; IDA.  
 DR GO; GO:0003680; P:AT DNA binding activity; IDA.  
 DR GO; GO:0000217; F:DNA secondary structure binding activity; IDA.  
 DR GO; GO:0003690; F:double-stranded DNA binding activity; IDA.  
 DR GO; GO:0007076; P:mitotic chromosome condensation; IMP.  
 DR InterPro; IPR003405; SMC\_C.  
 DR InterPro; IPR003395; SMC\_N.  
 DR Pfam; PF02483; SMC\_C; 1.  
 DR Pfam; PF02463; SMC\_N; 1.  
 DR Prodom; P000006; ABC transporter; 1.  
 KW DNA condensation; Mitosis; Cell cycle; ATP-binding; Coiled coil;  
 KW Nuclear protein.  
 KW NP\_BIND .32 .39 ATP (POTENTIAL).  
 FT DOMAIN 172 469 COILED COIL (POTENTIAL).  
 FT DOMAIN 470 677 FLEXIBLE HINGE.  
 FT DOMAIN 678 1027 COILED COIL (POTENTIAL).  
 FT DOMAIN 1084 1119 ALA/ASP-RICH (DA-BOX).  
 SQ SEQUENCE 1170 AA; 133927 MW; 14281AAE109621F CRC64;

Query Match 5.3%; Score 120; DB 1; Length 1170;  
 Best Local Similarity 18.6%; Pred. No. 7.5;  
 Matches 108; Conservative 82; Mismatches 170; Indels 220; Gaps 22;

3 KMIRPFALEYSKSGKTQENRNGSIGPSIVCKSIQNNQENSLQEOGEPDLTVNRQOE 62  
 261 KMLNEIFV-----KTSSEE-----ISLNEDEVEIKQKEKEKEKEKTSKLEN 303  
 63 QNTQGGDGL-DLSTKTSI-----KSEESICDPSE----- 94  
 304 KE---NGLNLEISRLKTSLSIKVENLNDTEKSKALESEIASSSKLIKKSAVANTEK 359  
 95 --NSVAGRLHNRREDYVER-----SAEPADG---LISKALDIOGSLDINX 136  
 360 DYKMQEOELSKORDLYKKEBELVLTITGISGTGAADGYNAQLAKAKTELVEVSLATK 419  
 137 AGILYGIPOKTLHLLEALPAPKPSFKXKTRDPHDSYVSKSKETCAVLQVALMARAO 196  
 420 SSMKMEELKKELLT-IE--PKLKEATKQNELNVKH---VKQCGETCDKLR-----ARLV 467  
 197 AERTKSKNLLETSEIKPPT--ASTYHLQHLQKKVTO----- 233  
 468 EYGFPSRLKDKOREDKLKHSHYQTCNKSEYLRKRVTLLEPNYTKPYENFEASEVHGAV 527  
 234 ---FKENKSLQYERSNPTVQAKIPQLRVSVSKSQPDGSGLLDVWVYQSKTSS--VLEGS 289  
 528 GQLFDINDNINRYALALOTCA-----GGRLPNVVVDSDIATOLEBNG 570  
 290 ALQKLNLPKQ-----SYFLGDLSPCLNSKNGVNDGTSNTTEGL--DKKDSQ 354  
 571 RLKREKVTIIPDKIYTRPISQVLDLAKKIAPGVKELAINLRPDESITKAMEFIKGNL 630  
 305 ECGSGVTSSVD-----SYFLGDLSPCLNSKNGVNDGTSNTTEGL--DKKDSQ 354  
 631 ICEDEFTAKKITFHPKIRARSITLQGD-----VYDEGTLGSGSRNTSLSLLVDIÖKYNQ 685  
 355 PKKRGYROYDHEIMEEAIAMWMSGKMSVSKAQGIYGVPHSTLEKVERSGTLKTPK 414  
 666 IOKIETIQADLNHTEEL-----QTOYATSOQTKITQSDLN 722  
 415 KKLRLPDTGLVMTDSGT-----GSCNNSK 440  
 723 LSLHKLIDLAKRNLDANPSSQIARNEELDRDIGECNEIK 762

## RESULT 9

HS68\_DROME STANDARD; PRT; 635 AA.  
 AC 097125;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heat shock protein 68.  
 GN HSP68 OR CG5436.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_taxid=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA McColl G., McKechnie S.W.;  
 RT "The heat shock gene hsp68 of D. melanogaster";  
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chang M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borrova D., Botchan M.R., Bouck C., Broststein P., Brothier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chertis J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milbina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Releart C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svateks R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Gibbs R.A., Myers E.N., Rubin G.M., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF096275; AAD16140.1; -  
 DR HSPB: AE003746; AAF56230.1; -  
 DR HSPB: P19120; 3HC.  
 DR FlyBase; FBgn001230; Hsp68.  
 DR InterPro; IPR001023; Hsp70.  
 DR Pfam; PF00012; HSP70; 1.  
 DR PRINTS; PR00301; HEATSHOCK70.  
 DR ProDom; PD000089; Hsp70; 1.  
 DR PROSITE; PS00297; HSP70.1; 1.  
 DR PROSITE; PS00329; HSP70.2; 1.  
 DR PROSITE; PS01036; HSP70.3; FALSE NEG.  
 DR ATP-binding; Heat shock.  
 KW SEQUENCE 635 AA; 69743 MW; B3A429D415BA8035 CRC64;

Query Match 5.2%; Score 118; DB 1; Length 635;  
 Best Local Similarity 21.0%; Pred. No. 4.4;  
 Matches 85; Conservative 58; Mismatches 152; Indels 110; Gaps 18;

55 LTVNRMOEQNT---QQGDGVLDLSTKRTSIKSEESSICDPSESVAGRLHRNEDYVER 111  
 182 LDKKLGKGRNVLPDLGGGTFDVS---LTIDESSLE---VNSTADDTLGGEDPUNR 234  
 112 SA-EPADGLSKALKDIOGALDINKAGILYIPQKL-----LHLALPAGKPAFK 164  
 235 LVNHFAEERKRYKKDLRSNPRALRLRTAERAKRTLSSTSEASLEIDALYEG----- 288  
 165 NKTDPHNSYSYKSKETCAVLQKALMARQAERTKS-KLNLSTSEIKFPPTASTYLH 223  
 289 ---HDFSVKVRARFEELCGDLFRNTL-----EPVEKALDAKDKSQI-----H 330  
 224 QLTQKMTQPKENKESLQYETSNPTVOLKI POLRVS SVSKSOPDGLDMYOVSKTS 283  
 331 DIVLVGSTRPKVONLQNFPGKTLWLT-----NDEBA---VAAGAAQQA 375  
 284 SVLEGSALQKLNILPKONKIECGPVTHSSVSYFLHGDLSPLCLNSKNGTDSSENT 343  
 376 AILSGDKSEIKDVLV-----DVAFLSLGIE-----TAGCV 407  
 344 EDGDKRKSQPKRKGROYDHEIMEEATAM-VMSGRMSVSKAGIYGVPHSTLEKV 402  
 408 MTKLIERSRIPCKQSKTFTTYADN-OPAVTIYQFEERALTJKNVNLGTFDILT----- 460  
 403 KERSGTLPKPKK-----LRLPDGLVNMV--DSGTSGCKN 437  
 461 -----GVPPAPRGVPRKIDVTFDLDANGILNVTKEGCTGNAKN 498

RESULT 10  
 RFCL\_HUMAN STANDARD; PRT; 1147 AA.  
 ID\_RFCL\_HUMAN AC P35251;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Activator 1 140 kDa subunit (Replication factor C large subunit) (A1 140 kDa subunit) (RF-C 140 kDa subunit) (Activator 1 large subunit)  
 DE (DNA-binding protein PO-GA).  
 GN RFCL OR RFCL140.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 469-480, 571-580 AND 677-699.  
 RX MEDLINE=94068535; PubMed=8248204;  
 RA Bunz F., Kobayashi R., Stillman B.;  
 RT "CDNAs encoding the large subunit of human replication factor C";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11014-11018(1993).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93290676; PubMed=8512577;  
 RA Lu Y., Zelt A.S., Riegel A.T.;  
 RT "Cloning and expression of a novel human DNA binding protein, PO-GA";

RL Biochem. Biophys. Res. Commun. 193:779-786(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hepatoma;  
 RA Rajavashisth T.B., Tripathi S.;  
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP FUNCTION, AND INTERACTION WITH PCNA.  
 RX MEDLINE=97153138; PubMed=8998859;  
 RA Mossi R., Jonsson Z.O., Allen B.L., Hardin S.H., Huebner U.;  
 RT "Replication factor C interacts with the C-terminal side of  
 RT proliferating cell nuclear antigen";  
 RL J. Biol. Chem. 272:1769-1776(1997).  
 RN [5]  
 RP DNA-BINDING ACTIVITY.  
 RX MEDLINE=96371221; PubMed=9705493;  
 RA Allen B.L., Uhlmann F., Gaur L.K., Mulder B.A., Posey K.L.,  
 RA Jones L.B., Hardin S.H.;  
 RT "DNA recognition properties of the N-terminal DNA binding domain  
 RT within the large subunit of replication factor C";  
 RL Nucleic Acids Res. 26:3877-3882(1998).  
 CC -I- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE  
 CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS  
 CC PCNA AND ACTIVATOR 1. THE 140 SUBUNIT BINDS TO THE PRIMER-TEMPLATE  
 CC JUNCTION. BINDS THE PO-B TRANSCRIPTION ELEMENT AS WELL AS OTHER  
 CC GA RICH DNA SEQUENCES. COULD PLAY A ROLE IN DNA TRANSCRIPTION  
 CC REGULATION AS WELL AS DNA REPLICATION AND/OR REPAIR. CAN BIND  
 CC SINGLE- OR DOUBLE-STRANDED DNA.  
 CC -I- FUNCTION: INTERACTS WITH C-TERMINUS OF PCNA. 5' PHOSPHATE RESIDUE  
 CC IS REQUIRED FOR BINDING OF THE N-TERMINAL DNA-BINDING DOMAIN TO  
 CC DUPLEX DNA, SUGGESTING A ROLE IN RECOGNITION OF NON-PRIMER  
 CC TEMPLATE DNA STRUCTURES DURING REPLICATION AND/OR REPAIR.  
 CC -I- SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND  
 CC 36.5 KDA THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP.  
 CC -I- SUBCELLULAR LOCATION: Nuclear.  
 CC -I- TISSUE SPECIFICITY: WIDE TISSUE DISTRIBUTION. UNDETECTABLE IN  
 CC PLACENTAL TISSUE.  
 CC -I- SIMILARITY: BELONGS TO THE ACTIVATOR 1 140 KDA SUBUNIT FAMILY.  
 CC -I- SIMILARITY: Contains 1 BRCT domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L14922; AAA86853.1; -  
 DR EMBL: L23320; AAA16121.1; -  
 DR EMBL: Z22642; CAA80355.1; -  
 DR EMBL: AF040250; AAB99788.1; -  
 DR PIR; A49651; A49651.  
 DR PIR; JN0599; JN0599.  
 DR Genew; HGNC:9969; RFCL.  
 DR GK: P35251; -  
 DR MIM; 102579; -  
 DR GO; GO:0005663; C:DNA replication factor C complex; TAS.  
 DR GO; GO:0005524; F:ATP binding activity; TAS.  
 DR GO; GO:0008047; F:enzyme activator activity; TAS.  
 DR GO; GO:0006261; P:DNA dependent DNA replication; TAS.  
 DR GO; GO:0007004; P:telomerase-dependent telomere maintenance; TAS.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003959; AAA\_ATPase\_centre.  
 DR InterPro; IPR001357; BRCT.  
 DR Pfam; PF00004; AAA; 1.  
 DR Pfam; PF00004; AAA; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS01072; BRCT; 1.  
 KW DNA replication; ATP-binding; Transcription regulation; DNA-binding;  
 KW Activator; Nuclear protein; Zinc-finger; Polymorphism;

```

FT DOMAIN 402 480 BRCT.
FT NP BIND 650 657 ATP (POTENTIAL).
FT 7N FING 749 766 C2HC-TYPE (POTENTIAL).
FT DOMAIN 1120 1124 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT VARIANT 598 598 I -> V (IN DBSNP:2066791).
FT FTID=VAR_014860.
FT CONFLICT 326 326 E -> K (IN REF. 1).
FT CONFLICT 613 613 L -> R (IN REF. 1).
FT CONFLICT 629 629 S -> A (IN REF. 1).
FT CONFLICT 640 640 N -> G (IN REF. 1).
FT CONFLICT 676 676 R -> A (IN REF. 1).
FT CONFLICT 1075 1075 A -> S (IN REF. 1).
SQ SEQUENCE 1147 AA; 128282 MW; 58C2878FDD2496D9 CR64;

Query Match
Best Local Similarity 21.5%; Score 118; DB 1; Length 1147;
Matches 109; Conservative 70; Mismatches 167; Indels 160; Gaps 25;

QY 3 KMIROPALEYISK-----SGKTQEN---RNGSIGPSIVCKSIOMNOAENSIOE 47
DB 97 KISRQDPVTYISETDEBDDFMCKKAASKENGRSTNSLGT-----NMKNIEVTKTKN 152
QY 48 EOEGLDLTVNRMOQNTQOGDGLDLSTYK--TSIKSESSICDPS--SENSVAGRLHR 103
DB 153 KPLSPDKLTPSTVLN---YFGTGSVQSRNKNVASKRKELSONTDESGINDAIAKQQL 209
QY 104 NREDVER---SAPFADGL-----LSKALKDIQSG-----ALDINKAGILVGIPO 145
DB 210 DEDAELEQOLHEDEFPARTLMDPEPKTKARKOTEAGETFSVOANLSKA-----E 262
QY 146 KTLHLLEALPAGKPASFPNKTREDFHD---SYS-----YKDSKETCAVLQVALLMARAO 196
DB 263 KKHVYH-----KVKTAQVSDERKSYSPKQSKYTESKSS----- 296
QY 197 AERTKSKLNLETSEIKFPASTYVHLQI-----TLQWQVTFKEKNESIQYETSN 247
DB 297 ---OQHSKSAADKIEVSSPKASKLAIWKRESSYKIEBPVASKRENAILKGETTKT 353
QY 248 PTVQAKIQQLRVSSV-----SKQPDGS-----GLDLY 275
DB 354 PKKTSPPKKESSVPESEKRTNYQAVSYLNRGPRALSKSEI PKGAENCLGLLFV 413
QY 276 MYQVSKTSSVLEGAQLQKLNLPKONKIEGSGPYTH--SSVDSYFLHGDLSPCLNSKN 333
DB 414 I-----TGVLESIERDEKSLIERX-----CGKTTGAVSKKTYLVWGRDSGSKSDXA 462
QY 334 GTVDGTSNTEDGDLDRKDSKOPRKRGROYDHEIMEBALMVNSGKMSVSKAOGIYGV 393
DB 463 AAL-CTKIIDEDGLNLIRTMGKKS-----KYELAVET-EMKESKLERTPQKVOG- 513
QY 394 PHSTLEKYV---KERSGTLKTPPKK 416
DB 514 ---KRTKISPKSESKSKSRPTSK 534

RESULT 11
KMAH_DICDI STANDARD; PRT; 1146 AA.
AC P42527;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myosin heavy chain kinase A (EC 2.7.1.129) (MHCK A).
GN MHCKA OR MHCKA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=AX3;
RX MEDLINE=9512486; PubMed=7822274;
RA Fuley L.M., Medley Q.G., Cote G.P., Egelhoff T.T.;
"Structural analysis of myosin heavy chain kinase A from

```

```

RT Dictyostelium. Evidence for a highly divergent protein kinase domain,
RT an amino-terminal coiled-coil domain, and a domain homologous to the
RT beta-subunit of heterotrimeric G proteins."
RL J. Biol. Chem. 270:523-529(1995).
RN [2]
RP CHARACTERIZATION OF THE CATALYTIC DOMAIN.
RC STRAIN=AX3;
RX MEDLINE=97207233; PubMed=9054368;
RA Cote G.P., Luo X., Murphy M.B., Egelhoff T.T.;
RT "Mapping of the novel protein kinase catalytic domain of
RT Dictyostelium myosin II heavy chain kinase A."
RL J. Biol. Chem. 272:6846-6849(1997).
CC -! FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION
CC OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN
CC REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.
CC REQUIRES AUTOPHOSPHORYLATION FOR ACTIVITY.
CC -! CATALYTIC ACTIVITY: ATP + [myosin heavy-chain] = ADP + [myosin
CC heavy-chain] phosphate.
CC -! COFACTOR: MAGNESIUM OR MANGANESE.
CC -! SUBUNIT: Oligomer.
CC -! DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN WITH PROBABLE COILED COIL
CC STRUCTURE, A CENTRAL NONREPEATITIVE CATALYTIC DOMAIN, AND A C-
CC -! PTM: THE N-TERMINUS IS BLOCKED.
CC -! SIMILARITY: Contains 7 WD repeats.
CC -! SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U16856; AAA6070.1; -.
CC DR PIR: A55532; A55532.
CC DR DICTYDB: DD01086; mhka.
CC DR InterPro: IPR004166; MHCK_EF2_kinase.
CC DR Pfam: PF02816; Alpha_kinase_1.
CC DR Pfam: PF00400; WD40_7.
CC DR PRINTS: PR000320; GPROTEINBRPT.
CC DR ProDom: PD000018; WD40_2.
CC DR SMART: SM00320; WD40_7.
CC DR PROSITE: PS00678; WD_REPEATS_1; 4.
CC DR PROSITE: PS50082; WD_REPEATS_2; 5.
CC DR Transferrase: Serine/threonine-protein kinase; ATP-binding; Repeat;
CC WD repeat; Phosphorylation; Coiled coil.
CC KW DOMAIN 100 120
CC FT DOMAIN 144 148
CC FT DOMAIN 175 181
CC FT DOMAIN 187 241
CC FT DOMAIN 297 502
CC FT DOMAIN 345 348
CC FT DOMAIN 438 441
CC FT DOMAIN 500 551
CC FT
CC FT DOMAIN 552 852
CC FT NP BIND 778 783
CC FT REPEAT 867 897
CC FT REPEAT 910 938
CC FT REPEAT 952 980
CC FT REPEAT 993 1021
CC FT REPEAT 1033 1061
CC FT REPEAT 1073 1101
CC FT REPEAT 1114 1142
CC SQ SEQUENCE 1146 AA; 128945 MW; 98D83177948B573 CR64;

Query Match
Best Local Similarity 19.0%; Score 117; DB 1; Length 1146;
Matches 94; Conservative 82; Mismatches 182; Indels 136; Gaps 20;

```

```

OY 33 C K S I O M N A E -- N S I O E B O E G P L D L T V N R M O E --- N T O O G D G V L D L S T K T K S I K S E S S 87
DB 55 C S S F L V S A P A E P D N H K D D A Q P H L Q A V E K P D H Q D L H T Q --- L W A H F T E Q M E D O L E T K T M 110
OY 88 I C D P S S E N S V A G R L H R N R E D Y V E R S A E F --- A D G L S K A L K D I O S G A L D I N K 136
DB 111 K V N R N H T S L G N V O T K L D E G I E K M A F A K V E Q O O Q O A R L I T Q O I E K K S I S S P L V K 170
OY 137 A G I L Y G --- I P Q K T L I H E A L P A G K A S F K N K T R D P H D S Y S Y K D 178
DB 171 G G I S G G G S G G D S F D G A N I S M S T S K O E L O Q E L O S L --- S I K M K E L T E I S D E L S O K L 226
OY 179 S K E T A V Q I K V A L M A R A Q A E R T E K - S K I N L U --- E N S E I K E P T A S T Y L --- 222
DB 227 E R S T G N I D I K I --- K R I E G V E N E I K D K R O L V S T I D S I G K T D S I G Y L E S S I I K K V E E K 283
OY 223 --- H O L T L O K W A T O F E K --- N E S L O Y E T S N P V O L K I P O L --- R 258
DB 284 E K K S E Q O Q L L F D S K I E S L K O K I I E T Q O L D T S E V A R K L K L E S T S G N I M A G I N G T S G R 343
OY 259 V S V S K S O P D S G L --- D W Y Q V S K T S S V L E S A L O K L K N I L E K O N K I E C S G P V T 311
DB 344 P S S S H F I P S S V S A A N N I N K N E I M E B V K V E E K L O K K I R E I D N T K A E L S K V E R S V D N 403
OY 312 H S V S Y T L H G D L S P L C L N S K N G T V D G S E N T E D G L D R K D S --- K O P R K R G R Y --- 362
DB 404 R S E I E G --- L E K D C K N O P D - K O D N K I K O V E D L K K S D L L M O N N L K K Y N E F V D R E 456
OY 363 --- R O Y D H E I M E - E A I A M W S G M S V S K A Q I Y G V P H S T I L E 399
DB 457 R D R E S E R L K O D S I K R L E Q N K K I E A L I Q E G N E Q V E R L R E A S I S P --- I S S V P K S P I - 512
OY 400 Y K V E R S G T L K T P P 413
DB 513 - T T R K S S I I L N S P P 525

RESULT 12
HKR1 YEAST STANDARD; PRT; 1802 AA.
ID HKR1 YEAST STANDARD; PRT; 1802 AA.
AC P41809;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hasegawa MRAKII killer toxin-resistant protein 1 precursor.
GN HKR1 OR YDR420W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YNN 295;
RA MEDLINE=94156857; PubMed=8113191;
RA Kasahara S., Yamada H., Mio T., Shitatori Y., Miyamoto C.,
RA Yabe T., Nakajima T., Ichishima E., Furuchi Y.;
RT "Cloning of the Saccharomyces cerevisiae gene whose overexpression
RT overcomes the effects of Hm-1 killer toxin, which inhibits
RT beta-glucan synthesis."
RL J. Bacteriol. 176:1488-1499 (1994).
CC - FUNCTION: COULD REGULATE BETA-GLUCAN SYNTHESIS. OVEREXPRESSION
CC - PROVIDES RESISTANCE TO HM-1 KILLER TOXIN.
CC - SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC - PTM: COULD BE O-GLYCOSYLATED IN SERINE/THREONINE RICH DOMAIN.
CC - SIMILARITY: SOME, TO YEAST MSB2.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC EMBL, S69101; AB30051.1; -.
DR SGD; S0002828; HKR1.
KW Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 1802
FT TRANSMEM 1486 1506
FT DOMAIN 23 1478
FT 453 788
FT REPEAT 453 480
FT REPEAT 481 508
FT REPEAT 509 536
FT REPEAT 537 564
FT REPEAT 565 592
FT REPEAT 593 620
FT REPEAT 621 648
FT REPEAT 649 676
FT REPEAT 677 704
FT REPEAT 705 732
FT REPEAT 733 760
FT REPEAT 761 788
FT CARBOHYD 24 24
FT CARBOHYD 1252 1252
FT CARBOHYD 1293 1293
FT CARBOHYD 1342 1342
FT CARBOHYD 1400 1400
SQ SEQUENCE 1802 AA; 188890 MW; E344CA6469785A24 CRC64;

Query Match 5.2%; Score 117; DB 1; Length 1802;
Best Local Similarity 19.6%; Pred. No. 20;
Matches 106; Conservative 82; Mismatches 173; Indels 180; Gaps 25;

OY 35 S I O M N A E N S I O E B O E G P L D L T V N R M O E --- N T O O G D G V L D L S T K T K S I K S E S S 78
DB 956 S A K I S S I O S L O S S T K P Y P - T A N K N T E T S G R S T V N S F L Y T S A A P D N E K F S A T P T E I 1014
OY 79 T S I K S E E S - S I C D P S S E N S V A G R L H R N R E D Y V E R S A E F A D G L S K A L K D I O S G A L D I N K 136
DB 1015 T T I S S S H A V S L S I P S S H N S V T G L S H - N F V D S K S A T S F - G Y S S S I S S I K --- 1063
OY 137 A G I L Y G I P O K T L I H E A L P A G K A S F K N K T R D P H D S Y S K D S E T C A V I O K A L M A R A Q 196
DB 1064 --- L S K E T I P A S K S V S --- N T Q E R I T S F T --- S T L R A N S Q 1094
OY 197 A E R T E - S K I N L I E T S E I K F - P T A S - T Y L H O L T L O Q M V --- T O F K 235
DB 1095 S E K E G N S V G S L O S S H I S S N P S I S T N T K V D S K S L S R K V K T M G E N G E T G L T T K T O Y K 1154
OY 236 E K N E --- S L O Y E T S --- 246
DB 1155 S S S E T S G S Y S R S F T K I S I G P A T T A V Q T A S T N S V F T A P A L S T Y P T P P S P N S Y A M L P T A 1214
OY 247 --- N E T V O L K I P O L A R V S S V S K O P G S G L D M V Q V S T S S V L E S G A L 291
DB 1215 I I V E S S E T G P T T A S F N S I T G L P A I E P A V A A S E P I N H T L I T I G F A A N Y V F L V O N P L 1274
OY 292 Q - K L K N I L P K O N K I E C S G P V T H S V D S Y F L H G D L S P L C L N S K N G V D G T - S E N T E D G L D 348
DB 1275 S S A Q I F P L P L V L K Y P P S N - T S E L D N S I - G E L S T F I L S Y S G S S T T L S P K S I S L S 1330
OY 349 - R K D S K O P R K R G R Y Q Y D H E --- I M E B A I M W S G M S V S K A Q I Y G V P H S T - 397
DB 1331 V V K K K N Q O K N A T K S E D L H P P O V D T S I A V K I V M V D S K A Y I V S A V E V F P T A V T 1390
OY 398 - L E Y K V E R S G T L K T P K K L R --- L P D T G --- L Y N M D S G --- T O S C K N S S 439
DB 1391 Y L O O L I D E N S T Y S N Q T P L R S L A G I D S G I P L G U L T I L Y G S G D G G V P S L T S S S V L D S S 1450
OY 440 K 440

```

Db 1451 K 1451

RESULT 13

MYH3\_HUMAN STANDARD; PRT; 1940 AA.

AC P11055; Q15492;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic  
MYH3.  
GN MYH3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89263803; PubMed=2726495;  
RA Ellier M.S., Steedman H.H., Sylvester J.E., Ferteis S.H., Wu Q.-L.,  
RA Rubinstein N.A., Kelly A.M., Sarkar S.;  
RT "Nucleotide sequence of full length human embryonic myosin heavy  
chain cDNA.";  
RL Nucleic Acids Res. 17:3591-3592(1989).  
RN [2]  
RP SEQUENCE OF 774-1940 FROM N.A.  
RX MEDLINE=90033298; PubMed=2806546;  
RA Ellier M.S., Steedman H.H., Sylvester J.E., Ferteis S.H., Wu Q.-L.,  
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;  
RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence  
conservation of the myosin rod, chromosomal locus and isoform  
specific transcription of the gene.";  
RL FEBS Lett. 256:21-28(1989).  
RN [3]  
RP SEQUENCE OF 856-1940 FROM N.A.  
RX TISSUE=skeletal muscle;  
RC MEDLINE=90235862; PubMed=1691980;  
RA Bobber E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,  
RA Arnold H.H.;  
RT "Identification of three developmentally controlled isoforms of human  
myosin heavy chains.";  
RL Eur. J. Biochem. 189:55-65(1990).  
RN [4]  
RP SEQUENCE OF 856-1940 FROM N.A.  
RX MEDLINE=8936648; PubMed=2771643;  
RA Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;  
RT "Expression and DNA sequence analysis of a human embryonic skeletal  
muscle myosin heavy chain gene.";  
RL Nucleic Acids Res. 17:6167-6179(1989).  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE  
AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL  
MUSCLE.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
SUBFRAGMENT (S2).  
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
CC -1- SIMILARITY: Contains 1 IQ domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL, X13988; CAA32167.1; -  
CC EMBL, X13100; CAA31492.1; -  
CC EMBL, X51593; CAA35942.1; -  
CC EMBL, X15696; CAA33731.1; -  
CC PIR, S04090; S04090.  
CC DR HSSP, P13538; 2MYS.  
CC DR GeneW, HGNC:7573; MYH3.  
CC DR MIM, 160720; -  
CC DR GO, GO:0007517; P:muscle development; TAS.  
CC DR InterPro, IPR000048; IQ\_region.  
CC DR InterPro, IPR001609; myosin\_head.  
CC DR InterPro, IPR004009; myosin\_N.  
CC DR InterPro, IPR002928; Myosin\_tail.  
CC DR Pfam, PF00612; IQ\_2.  
CC DR Pfam, PF00613; myosin\_head; 1.  
CC DR Pfam, PF02736; Myosin\_N; 1.  
CC DR Pfam, PF01576; Myosin\_tail; 1.  
CC DR PRINTS, PR00193; MYOSINHEAVY.  
CC DR ProDom, PD000355; myosin\_head; 1.  
CC DR SMART, SM00242; MYSC1.  
CC DR SMART, PS50096; IQ\_1.  
CC DR ProSITE, PS50096; IQ\_1.  
CC KW Myosin; muscle protein; Coiled coil; Thick filament; Actin-binding;  
CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
CC Multigene family.  
CC FT DOMAIN 1 781 MYOSIN HEAD-LIKE.  
FT DOMAIN 782 811 IQ.  
FT NP\_BIND 840 1933 COILED COIL (POTENTIAL).  
FT FT 179 186 ATP (POTENTIAL).  
FT FT 656 678 ACTIN-BINDING.  
FT FT DOMAIN 758 772 ACTIN-BINDING.  
FT FT MOD\_RES 130 130 METHYLATION (SH-1) (POTENTIAL).  
FT FT MOD\_RES 696 696 ALKYLATION (SH-2).  
FT FT MOD\_RES 706 706 ALKYLATION (SH-1).  
FT FT CONFLICT 1331 1331 A -> G (IN REF. 3).  
FT FT CONFLICT 1391 1392 KK -> QE (IN REF. 1 AND 2).  
FT FT CONFLICT 1608 1609 SR -> RA (IN REF. 3).  
FT FT CONFLICT 1663 1664 RG -> QT (IN REF. 2).  
SQ SEQUENCE 1940 AA; 224035 MW; 43CA586CA4BA1253 CRC64;

Query Match 5.2%; Score 117; DB 1; Length 1940;  
Best Local Similarity 18.0%; Pred. No. 23;  
Matches 87; Conservative 84; Mismatches 183; Indels 130; Gaps 18;

QY 18 KTOEHRNRSIGPSIVCKSIQWNAQNSIQEEDGFLDTVRMGOENTQGD-----GV 71  
DB 1012 QAEEDKVNLSNKTYSKLEQVEDESSLEQEKRLVDLERNRK---LEGDKLTAOSI 1067  
QY 72 LDLSTFKTS-----IKSESSICDPSS-----EN 95  
DB 1068 LDLENDKOQDLERLKKQPFYCOLOSQVEDCTGLQFOQKIKELQAAIEELFEIEBAER 1127  
QY 96 SVAGRLHNRREDYVERSAEPAFDGLSKALKD---IOSGALDINKAGILYGIPOKTLTLHL 152  
DB 1128 ATRAKTEKORSQDYARELEE-----LSERLEAGVGTSTQIENLKK----- 1167  
QY 153 EALPAKGRPAKTRDPHDY-----SYDSKETCAVLQKVALMARQ 196  
DB 1168 -----RAAEFLKRLRDLAEATLQHEAMVATLRKKHADVVELGQIDNLRQVK--OKLE 1219  
QY 197 AERTE-KSKUNLLETSEIKPTASTYTLHQL--TLQKWVTOPEKKESSLOVETNSPTVOLK 253  
DB 1220 KEKSEFKLEIDLSSMSVSKSKANLEKICRTLEDQJSEARGKKEIORSISELTTOXS 1279  
QY 254 IPQLRVSSVSKSQDPGSLDVMVQVSVSTSVLEGSALQKKNLIPKONKIE--CSGPVT 311  
DB 1280 RLQTAGELSLQLEEKESIVS---QLSRSKQAFV-QQTEELKQPLEENKAKNALAHALQ 1335

QY 312 HSSVDSYFLH-----GDLSPCLNSKNGTVDGTSNTEDGLDRKDSKOPKCKGR 361  
 DB 1336 SSRHDCDLREQEEBEGSKLQALSKANSEVAMQMTKYETDIAIQTEBELKAKKLLA 1395  
 QY 332 YROYHEIMEEAI-AMVMSGKMSVSKAGCIGVPHSTLEYKYK-----ESGTLTKTPK 414  
 DB 1396 ORLOPSEOEAVNAKCAKSLKTKORLOG-----EVEDIMVDVERANSILAAALD 1444  
 QY 415 KKLK 418  
 DB 1445 KKOR 1448  
 RESULT 14  
 POLG\_EC23C  
 ID\_POLG\_EC23C STANDARD; PRT; 2188 AA.  
 AC 09YID8;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polypeptide [contains: Coat protein VP0 (PIAB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein 2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPg (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 DE Picornavirus 23 (strain CTR6-6760) (P3D).  
 OS Echovirus 23 (strain CTR6-6760) (Human parechovirus 2).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Parechovirus  
 NC NCB1\_TaxID=122961;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98454792; PubMed=9783471;  
 RX Oberste M.S., Maher K., Pallansch M.A.;  
 RT "Complete sequence of echovirus 23 and its relationship to echovirus 22 and other human enteroviruses."  
 RL Virus Res. 56:217-223 (1998).  
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT IS A CYSTEINE PROTEASE.  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polypeptide. In other picornavirus reactions Gln may be substituted for Gln, and Ser or Thr for Gly.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + (RNA) (N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP0, VP1, AND VP3.  
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD NATURE PROTEINS.  
 CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL, AF055846; AAC79756.1; -.  
 DR MEROPS; C03.023; -.  
 DR InterPro; IPR004004; Calict\_pol\_hel.  
 DR InterPro; IPR007053; NC.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF04970; NC; 1.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICIVIRUSNS.  
 KW Polypeptide; Coat protein; Core protein; Transferase;

KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 FT CHAIN 1 290  
 FT CHAIN 291 549  
 FT CHAIN 550 784  
 FT CHAIN 785 931  
 FT CHAIN 932 1053  
 FT CHAIN 1054 1382  
 FT CHAIN 1383 1499  
 FT CHAIN 1500 1519  
 FT CHAIN 1520 1719  
 FT CHAIN 1720 2188  
 FT SITE 772 774  
 FT ACT\_SITE 1678 1678  
 FT ACT\_SITE 1696 1696  
 SQ SEQUENCE 2188 AA; 246602 MW; 02CC77D0A5ED3D93 CRC64;  
 Query Match 5.2%; Score 117; DB 1; Length 2188;  
 Best Local Similarity 22.5%; Pred. No. 26;  
 Matches 87; Conservative 61; Mismatches 150; Indels 88; Gaps 19;  
 QY 70 GULDSTKTSIKSESSICDPSSNSVAGR-----LHRNREYVERSAEF--A 116  
 DB 1323 GKLVSQAMSTMSYGE--CWEVSKN--GRDWETLKLKDLVQKITEDYERQKNVAMK 1376  
 QY 117 DGLSKLKDIQSGALINKAGILYGP-----QKTLHLLEALPKGKPSFKNKT 167  
 DB 1377 QOLENQTLDDDD-AVSYIKNFPDAIPYIDEVINIMSTLIEOMEAFIEPRSEVFK-- 1432  
 QY 168 RDFHDSYVSKDSKETCAVLQKVALARAQAERTKSKLNLETSEIFPTASTYLOTL 227  
 DB 1333 -CFAVLKPCHKGK-----QPKLMAGSACK-IKSLNLFIERKAKLVATVSAATSAISI 1483  
 QY 228 QKMTQREKNESIQTSNPTVOLKIPOLRVSSVSKSGPDGSLDVMYQVSKTSVLE 287  
 DB 1484 LLLVTKLFFKEESDERAYNPFLPITKPK-----GTFEVSQREFNGEAPYD 1529  
 QY 288 GSAQLQKNILPKNKI--ECSGPVTHSS--VDSYFLHG-----DLSPCLNSKNGT 335  
 DB 1330 G----QLEHITSQAAVYTGSTGTHLHCAGYQHDEITLHGHSIVYLEQOEDTLTHYKNKV 1585  
 QY 336 VDGTSN--TEDGLDRK-----DSKOPRKKGRYROYHEIMEBAIMVMSGKMSV 385  
 DB 1586 F--FIENPSYQVTLGSKPMDIALTKCKLPFRFKSKSKYTNKIGTSMILMWTGQGIT 1643  
 QY 386 KAQCIGVPHSTLEYKYERSGTUKT 411  
 DB 1644 KE--VQRVHSHG--GIKTREGTEST 1664  
 RESULT 15  
 NEKI\_HUMAN  
 ID\_NEKI\_HUMAN STANDARD; PRT; 1258 AA.  
 AC 096PY6; Q9Y594;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serine/threonine-protein kinase NEKI (EC 2.7.1.37) (NimA-related protein kinase 1) (NY-REN-55 antigen).  
 GN NEKI OR KIAA1901.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 CC NCB1\_TaxID=9606;  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=21456161; PubMed=11572484;  
 RA Nagase T., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XXI. The complete sequences of 60 new cDNA clones from brain which code for large proteins."  
 RT

```

NL DNA Res. 8:179-187(2001).
RL [2]
RP SEQUENCE OF 444-1258 FROM N.A. (ISOFORM 2).
RC TISSUE=renal cell carcinoma;
RX MEDLINE=99438124; PubMed=10508479;
RA Scanlin M.J., Gordan J.D., Williamson B., Stockert E., Bender N.H.,
RA Jorgensen C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA Old L.J.;
RT "Antigens recognized by autologous antibody in patients with
RT renal-cell carcinoma.";
RL Int. J. Cancer 83:456-464(1999).
CC -I- FUNCTION: PHOSPHORYLATES SERINES AND THREONINES, BUT ALSO APPEARS
CC TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF
CC MEIOSIS (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96PY6-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96PY6-2; Sequence=VSP_004870;
CC Note=No experimental confirmation available;
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. NIMA
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by . and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB067488; BAB67794.1; ALT_INIT.
DR EMBL; AF15113; AAD42879.1; -.
DR Genew; HGNC:7744; NEK1.
DR GK; Q96PY6; -.
DR MIM; 604568; -.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD006001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Translerase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;
KW Tyrosine-protein kinase; Alternative splicing.
FT DOMAIN 4 258 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT VASPLIC 478 521 Missing (in isoform 2).
FT FTID=VSP_004870.
FT CONFLICT 1232 1232 G -> E (IN REF. 2).
FT SO SEQUENCE 1258 AA; 142828 MW; 339C4BFA5612530 CRC64;

Query Match 5.2%; Score 116; DB 1; Length 1258;
Best Local Similarity 18.8%; Pred. No. 15;
Matches 89; Conservative 88; Mismatches 174; Indels 122; Gaps 22;

OY 11 EYKSGKTOENRNGSIGPVSIVKSIQNMNAENSLQEEQEPGLDTVNRMQONT--QQG 68
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 536 EFLORKEBAMONKRAEBGHMYTLARLIPLON-FNEHQI KAKLRGEKKEANSHSEGG 594
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 69 DGVLDLSTKK-TSTKSESSICDPSPSSVAGRLRHNRDVEDVSAF-FAGGLSKALKD 126
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 595 SEBDMMRKKESTLSLAHNAAPAAVLKE-----QLERKKEKVEYERKKVQWEEHLVAKGVS 649
OY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Qy 127 IQ-SGALDINKKAGILYIGPOKTLULHLALBAGKPRASPKKTRBDFHDSYVSXOSKECAV 165
Db 650 SDVSPPLGOGHTG-----GSPSKOQMSV-----ISTSLAKKEVGDDSLTDRRESEE 699
Qy 186 LOKV-----ALWARAOERTEKSKLNLLETSEIKFPASTYLH-----QLTLOK 229
Db 699 MOQTNNAISSKREILRLNENIKAKOEDBKONLSDTFEIN-----VHEDAKHEKEK 751
Qy 230 MTOFKENKES-----LOYETSNPTVOLKI POLRVS SVSKOPDGS----- 270
Db 752 SVSDBRKKMEAGQULVPLDELTLDTSPSTHE-----RHTVEGVIKLBNPNSPRAMGKSP 807
Qy 271 --GLDVM--YOVSKTSVLESGAL-----OKTKNILEPKONKIEC-----S 307
Db 808 TOSVLILTEAELOTELLETTLNTTIRSEISPEGEKXFKPLIGEKKVOCSIEHINPSAIVD 867
Qy 308 GPHTHSVD-----SYFLHG-----DLSPLCLNSKNOTVGTJSENTEDGL----- 347
Db 868 SPVETKSPFSEASPOMSLKEGNLBEPDDDETEILLOPSTG-----NKDESUPCTITD 921
Qy 348 ----DKDSKOPRKKRGRYROYDHEIMEEAIAMWMSGMSKSAQGIYGVPHS 396
Db 922 WJISEKEKEKE--TQSAADRITIOENVSVDGVSIVDOLSDIHIEPGRINDSOHS 973

```

Search completed: October 28, 2003, 12:02:38  
Job time : 21.0727 secs

**THIS PAGE BLANK (USPTO)**

STIC-Biotech/C

106818

**From:** Christina  
**Sent:** Monday, October 27, 2003 1:35 PM  
**To:** Davis, Minh-Tam; STIC-Biotech/ChemLib  
**Subject:** RE: Rush search request for 10/016768

RECEIVED

OCT 27 2003

(STIC)

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Monday, October 27, 2003 1:31 PM  
**To:** Chan, Christina  
**Subject:** FW: Rush search request for 10/016768

Please add:

Please search in commercial database, PDPUB, issued patent files and interferences:  
The polypeptide of SEQ ID NO:1.  
Thank you.

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Monday, October 27, 2003 12:23 PM  
**To:** Chan, Christina  
**Subject:** Rush search request for 10/016768

Please search in commercial database, PDPUB, issued patent files and interferences:  
The polypeptide of SEQ ID NO:8  
Thank you.  
MINH TAM DAVIS  
ART UNIT 1642, ROOM 8A01, MB 8E12  
305-2008

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

Hanley, Susan

107 007

**From:** Davis, Minh-Tam  
**Sent:** Tuesday, October 28, 2003 4:26 PM  
**To:** Hanley, Susan  
**Subject:** 10/016768

---

Thanks for the search results

Could you also do a rush search for the polypeptide SEQ ID NO:10 in commercial database, PGPUB, issued patent files and interference?

Thanks

MINH TAM DAVIS

ART UNIT 1642, ROOM 8A01, MB 8E12

305-2008

**THIS PAGE BLANK (USPTO)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 106813

TO: Minh-Tam Davis  
Location: cm1/8a01/8e12  
Art Unit : 1642  
Tuesday, October 28, 2003

Case Serial Number: 10/015768

From : Susan Hanley  
Location: Biotech-Chem Library  
CM1 6B05  
Phone: 305-4053

susan.hanley@uspto.gov

### Search Notes



**THIS PAGE BLANK (USPTO)**

106 813

**From:** Chan, Christina  
**Sent:** Monday, October 27, 2003 2:50 PM  
**To:** Davis, Minh-Tam; STIC-Biotech/ChemLib  
**Subject:** RE: Rush search request for 10/016768

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644  
 308-3973  
 CM-1, 9B19

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Monday, October 27, 2003 2:13 PM  
**To:** Chan, Christina  
**Subject:** FW: Rush search request for 10/016768

Please add:  
 Compare SEQ ID NO:8 with SEQ ID NO:10 to determine percent identity.  
 Thank you

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Monday, October 27, 2003 1:31 PM  
**To:** Chan, Christina  
**Subject:** FW: Rush search request for 10/016768

Please add:  
 Please search in commercial database, PDPUB, issued patent files and interferences:  
 The polypeptide of SEQ ID NO:1.  
 Thank you.

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Monday, October 27, 2003 12:23 PM  
**To:** Chan, Christina  
**Subject:** Rush search request for 10/016768

Please search in commercial database, PDPUB, issued patent files and interferences:  
 The polypeptide of SEQ ID NO:8  
 Thank you.

MINH TAM DAVIS  
 ART UNIT 1642, ROOM 8A01, MB 8E12  
 305-2008

Searcher: \_\_\_\_\_  
 Phone: \_\_\_\_\_  
 Location: \_\_\_\_\_  
 Date Picked Up: \_\_\_\_\_  
 Date Completed: \_\_\_\_\_  
 Searcher Prep/Review: \_\_\_\_\_  
 Clerical: \_\_\_\_\_  
 Online time: \_\_\_\_\_

TYPE OF SEARCH:  
 NA Sequences: \_\_\_\_\_  
 AA Sequences: \_\_\_\_\_  
 Structures: \_\_\_\_\_  
 Bibliographic: \_\_\_\_\_  
 Litigation: \_\_\_\_\_  
 Full text: \_\_\_\_\_  
 Patent Family: \_\_\_\_\_  
 Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
 STN: \_\_\_\_\_  
 DIALOG: \_\_\_\_\_  
 Questel/Orbit: \_\_\_\_\_  
 DRLink: \_\_\_\_\_  
 Lexis/Nexis: \_\_\_\_\_  
 Sequence Sys.: \_\_\_\_\_  
 WWW/Internet: \_\_\_\_\_  
 Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



LDLTVNMQGONTQGGVLDLSTKTSIKSEESSICDPSSSENVAGRLHNRREDYVERSAEFADGLSKAL  
LDANVLHTLMLAAGAMPKLDLDTQVGDIFKGLLVANSGLINNEGL--NLISASQENSGNASLLLOQO  
KDIOGSLDLINKAGILVIGIPQKTLHLHEALPAGKPSAFKUKTRDFHDSYKSKSKETCAVLQKVALMARAO  
OHOOHHOOHHOOQOOQHHVAAAYRRRLPKSETPETNSSLDPNDA SEDPLKIPSPKVS GPASSSSLS PGLVY  
AERTREKSNLLETSEIKFTASTYLLHQLTIQKAVTOPEKKNESLOYETSNPTVQLKIPOLRVSSVSKSQPD  
GHHPLNNNSLSTISNNSHSSSHRNGSNRSPHSASPMLAAVAOQGYSA GNSLLTSSSSSIQKMMASNIQ  
GSGLLDVMYQVSKTSSYLEGSA LQKLNILPKONKIEGSGPVTHSSVDYFLHGDLSPLCLNSKNGTYDGT  
ROINEOGQGESLRNGNVSDCSSNNGSSSLGKPPSISVAKIIGTDTSRFGASPNLISQOHS AHHLT HQ  
ENTEDGLDRKDSKQPRKKRGYROYDHEIMEALAMWSGKMSVSKAOGIYGVPHSTLEYKVKERSGTLKTP  
QOOOLSAQEBALGKTRPRGRKYNVDRDSDLVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKERHLMRPK  
PKKLRLPDPTGLYMMTDSGTGSCNNSKPV  
REP KPPDPDLVGLTGPANKLQDLKLAGPHGSKLSNALKNONNOAAAAA  
- 750 760 770 780 790 800 810  
420 430 440 X  
820 830 840 850 860

Diversity in the mechanisms of neuronal cell death.

Yuan Junying; Lipinski Marta; Degterev Alexei

Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue,  
02115, Boston, MA, USA

Neuron (United States) Oct 9 2003, 40 (2) p401-13, ISSN 0896-6273

Journal Code: 8809320

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: In Process

Neurons may die as a normal physiological process during development or as a pathological process in diseases. The best-understood mechanism of neuronal cell death is **apoptosis**, which is regulated by an evolutionarily conserved cellular **pathway** that consists of the **caspase** family, the Bcl-2 family, and the adaptor protein Apaf-1. **Apoptosis**, however, may not be the only cellular mechanism that regulates neuronal cell death. Neuronal cell death may exhibit morphological features of autophagy or necrosis, which differ from that of the canonical **apoptosis**. This review evaluates the evidence supporting the existence of alternative mechanisms of neuronal cell death and proposes the possible existence of an evolutionarily conserved **pathway** of necrosis.

.... as a pathological process in diseases. The best-understood mechanism of neuronal cell death is **apoptosis**, which is regulated by an evolutionarily conserved cellular **pathway** that consists of the **caspase** family, the Bcl-2 family, and the adaptor protein Apaf-1. **Apoptosis**, however, may not be the only cellular mechanism that regulates neuronal cell death. Neuronal cell...

... may exhibit morphological features of autophagy or necrosis, which

**THIS PAGE BLANK (USPTO)**

? ds

Set	Items	Description
S1	72840	APOPTOSIS
S2	322867	REVIEW
S3	3192	S1 AND S2
S4	361153	MECHANISM
S5	500	S3 AND S4

? s pathway??

S6 280848 PATHWAY??

? s s3 and s6

3192 S3  
280848 S6

S7 1057 S3 AND S6

? s caspase??

S8 13312 CASPASE??

? s s7 and s8

1057 S7  
13312 S8

S9 180 S7 AND S8

? rd

...examined 50 records (50)

...examined 50 records (100)

...examined 50 records (150)

...completed examining records

S10 179 RD (unique items)

? t s10/3,k,ab/1-10

10/3,K,AB/1

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

15526993 22924607 PMID: 14563117

**THIS PAGE BLANK (USPTO)**

257960 22654947 PMID: 12769773

**Apoptosis** induced by topoisomerase inhibitors.

Sordet Olivier; Khan Qasim A; Kohn Kurt W; Pommier Yves  
Laboratory of Molecular Pharmacology, Center for Cancer Research,  
National Cancer Institute, NIH, Bethesda, Maryland 20892-4255, USA.

Curr Med Chem Anti-Canc Agents (Netherlands) Jul 2003, 3 (4) p271-90

, ISSN 1568-0118 Journal Code: 101123597

Document type: Journal Article; Review; Review, Academic

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Topoisomerase inhibitors are among the most efficient inducers of **apoptosis**. The main **pathways** leading from topoisomerase-mediated DNA damage to cell death involve activation of **caspases** in the cytoplasm by proapoptotic molecules released from mitochondria. In some cells, apoptotic response also involves the death receptor Fas (APO-1/CD95). The engagement of these apoptotic effector **pathways** is tightly controlled by upstream regulatory **pathways** that respond to DNA lesions-induced by topoisomerase inhibitors in cells undergoing **apoptosis**. These include the proapoptotic Chk2, c-Abl and SAPK/JNK **pathways**, the survival PI(3)kinase-Akt-dependent **pathway** and the transcription factors p53 and NF-kappaB. Initiation of cellular responses to DNA lesions-induced by topoisomerase inhibitors is ensured by the protein kinases DNA-PK, ATM and ATR, which bind to DNA breaks. These kinases commonly called "DNA sensors" mediate their effects (DNA repair, cell cycle arrest and/or **apoptosis**) by phosphorylating a large number of substrates, including several downstream kinases such as c-Abl and the checkpoint protein Chk2. c-Abl induces **apoptosis** by activating cell death **pathways** (e.g., SAPK, p53 and p73) and inhibiting cell survival **pathways** [e.g., PI(3)kinase]. The DNA-damage regulating kinase Chk2, in addition to its role in cell cycle arrest and/or DNA repair, can induce **apoptosis** by phosphorylation/activation of the promyelocytic leukemia (PML) protein and p53. Finally, we will **review** the recent observations that support a role for topoisomerases in chromatin fragmentation during the execution phase of **apoptosis**.

**THIS PAGE BLANK (USPTO)**

15271751 22761780 PMID: 12879973

Cerebellar granule cells as a model to study mechanisms of neuronal **apoptosis** or survival in vivo and in vitro.

Contestabile Antonio

Department of Biology, University of Bologna, Italy.  
acontest@alma.unibo.it

Cerebellum (England) Jan-Mar 2002, 1 (1) p41-55, ISSN 1473-4222

Journal Code: 101089443

Document type: Journal Article; Review; Review, Academic

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Granule cells of the cerebellum constitute the largest homogeneous neuronal population of mammalian brain. Due to their postnatal generation and the feasibility of well characterized primary in vitro cultures, cerebellar granule cells are a model of election for the study of cellular and molecular correlates of mechanisms of survival/**apoptosis** and neurodegeneration/neuroprotection. The present **review** mainly deals with recent data on mechanisms and factors promoting survival or apoptotic elimination of cerebellar granule neurons, with a particular focus on the molecular correlates at the level of gene expression and induction of cellular signal **pathways**. The in vivo development is first analysed with particular reference to the role played by several neurotrophic factors and by the NMDA subtype of glutamate receptor. Then, mechanisms of survival/**apoptosis** are examined in the model of primary in vitro cultures, where the role of neurotrophins acting on cerebellar granule cells is followed by the large deal of data coming from the paradigm of potassium/serum withdrawal. The role of some key genes of the Bcl family, of some kinase systems and of transcriptional factors is primarily highlighted. Furthermore, the involvement of mitochondria, free radicals and proteases of the **caspase** family is considered. Finally, the use of cerebellar granule neurons in primary culture to experimentally address the issue of neurodegeneration and pharmacological neuroprotection is considered, with some comments on models at the borderline between necrosis and **apoptosis**, such as the excitotoxic neuronal damage. The overlapping of cellular signal **pathways** activated in granule neurons by apparently unrelated stimuli, such as neurotrophins and neurotransmitters/neuromodulators is stressed to put into light the special 'trophic' role played by activity in neurons. Finally, the advantage of designing and performing conceptually equivalent experiments on cerebellar granule neurons during development in vivo and in vitro, is stressed. On the basis of the reviewed material, it is concluded that cerebellar granule neurons have acquired a special position in modern neuroscience as one of the most reliable models for the study of neural development, function and pathology.

Cerebellar granule cells as a model to study mechanisms of neuronal **apoptosis** or survival in vivo and in vitro.

... model of election for the study of cellular and molecular correlates of mechanisms of survival/**apoptosis** and neurodegeneration/neuroprotection. The present **review** mainly deals with recent data on mechanisms and factors promoting survival or apoptotic elimination of...

... on the molecular correlates at the level of gene expression and induction of cellular signal **pathways**. The in vivo development is first analysed with particular reference to the role played by several neurotrophic factors and by the NMDA subtype of glutamate receptor. Then, mechanisms of survival/**apoptosis** are examined in the model of primary in vitro cultures, where the role of neurotrophins...

... factors is primarily highlighted. Furthermore, the involvement of mitochondria, free radicals and proteases of the **caspase** family is considered. Finally, the use of cerebellar granule neurons in primary

**THIS PAGE BLANK (USPTO)**

culture to experimentally...

... pharmacological neuroprotection is considered, with some comments on models at the borderline between necrosis and **apoptosis**, such as the excitotoxic neuronal damage. The overlapping of cellular signal **pathways** activated in granule neurons by apparently unrelated stimuli, such as neurotrophins and neurotransmitters/neuromodulators is...

Descriptors: **Apoptosis**--genetics--GE; \*Cell Survival--genetics--GE;  
\*Cerebellar Cortex--growth and development--GD; \*Neurons--metabolism--ME

10/3,K,AB/8

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

15257960 22654947 PMID: 12769773

**Apoptosis** induced by topoisomerase inhibitors.

Sordet Olivier; Khan Qasim A; Kohn Kurt W; Pommier Yves

Laboratory of Molecular Pharmacology, Center for Cancer Research,  
National Cancer Institute, NIH, Bethesda, Maryland 20892-4255, USA.

Curr Med Chem Anti-Canc Agents (Netherlands) Jul 2003, 3 (4) p271-90

, ISSN 1568-0118 Journal Code: 101123597

Document type: Journal Article; Review; Review, Academic

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Topoisomerase inhibitors are among the most efficient inducers of **apoptosis**. The main **pathways** leading from topoisomerase-mediated DNA damage to cell death involve activation of **caspases** in the cytoplasm by proapoptotic molecules released from mitochondria. In some cells, apoptotic response also involves the death receptor Fas (APO-1/CD95). The engagement of these apoptotic effector **pathways** is tightly controlled by upstream regulatory **pathways** that respond to DNA lesions-induced by topoisomerase inhibitors in cells undergoing **apoptosis**. These include the proapoptotic Chk2, c-Abl and SAPK/JNK **pathways**, the survival PI(3)kinase-Akt-dependent **pathway** and the transcription factors p53 and NF-kappaB. Initiation

**THIS PAGE BLANK (USPTO)**

? ds

Set	Items	Description
S1	1324	ID1 OR ID(W)1
S2	1210	ID2 OR ID(W)2
S3	397	S1 AND S2
S4	1431609	ANTIBOD?
S5	49	S3 AND S4
S6	34	RD (unique items)
S7	18	S6 AND PY<=2000

? s seq

S8 31073 SEQ

? s s7 not s8

18 S7

31073 S8

S9 15 S7 NOT S8

? t s9/3,k,ab/1-15

9/3,K,AB/1 (Item 1 from file: 155)

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

11974386 99419278 PMID: 10487839

Id-1 and Id-2 are overexpressed in pancreatic cancer and in dysplastic lesions in chronic pancreatitis.

Maruyama H; Kleeff J; Wildi S; Friess H; Buchler M W; Israel M A; Korc M  
Division of Endocrinology, Department of Medicine, University of California, Irvine, USA.

American journal of pathology (UNITED STATES) Sep 1999, 155 (3)

p815-22, ISSN 0002-9440 Journal Code: 0370502

Contract/Grant No.: CA-40162; CA; NCI

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Id proteins antagonize basic helix-loop-helix proteins, inhibit differentiation, and enhance cell proliferation. In this study we compared the expression of Id-1, Id-2, and Id-3 in the normal pancreas, in pancreatic cancer, and in chronic pancreatitis (CP). Northern blot analysis demonstrated that all three Id mRNA species were expressed at high levels in pancreatic cancer samples by comparison with normal or CP samples. Pancreatic cancer cell lines frequently coexpressed all three Ids, exhibiting a good correlation between Id mRNA and protein levels, as determined by immunoblotting with highly specific anti-Id antibodies. Immunohistochemistry using these antibodies

demonstrated the presence of faint Id-1 and Id-2 immunostaining in pancreatic ductal cells in the normal pancreas, whereas Id-3 immunoreactivity ranged from weak to strong. In the cancer tissues, many of the cancer cells exhibited abundant Id-1, Id-2, and Id-3 immunoreactivity. Scoring on the basis of percentage of positive cells and intensity of immunostaining indicated that Id-1 and Id-2 were increased significantly in the cancer cells by comparison with the respective controls. Mild to moderate Id immunoreactivity was also seen in the ductal cells in the CP-like areas adjacent to these cells and in the ductal cells of small and interlobular ducts in CP. In contrast, in dysplastic and atypical papillary ducts in CP, Id-1 and Id-2 immunoreactivity was as significantly elevated as in the cancer cells. These findings suggest that increased Id expression may be associated with enhanced proliferative potential of pancreatic cancer cells and of proliferating or dysplastic ductal cells in CP.

Id-1 and Id-2 are overexpressed in pancreatic cancer and in dysplastic lesions in chronic pancreatitis.

**THIS PAGE BLANK (USPTO)**

Sep 1999,

... proteins, inhibit differentiation, and enhance cell proliferation. In this study we compared the expression of Id-1, Id-2, and Id-3 in the normal pancreas, in pancreatic cancer, and in chronic pancreatitis (CP...

...between Id mRNA and protein levels, as determined by immunoblotting with highly specific anti-Id antibodies. Immunohistochemistry using these antibodies demonstrated the presence of faint Id-1 and Id-2 immunostaining in pancreatic ductal cells in the normal pancreas, whereas Id-3 immunoreactivity ranged from weak to strong. In the cancer tissues, many of the cancer cells exhibited abundant Id-1, Id-2, and Id-3 immunoreactivity. Scoring on the basis of percentage of positive cells and intensity of immunostaining indicated that Id-1 and Id-2 were increased significantly in the cancer cells by comparison with the respective controls. Mild to...

... and interlobular ducts in CP. In contrast, in dysplastic and atypical papillary ducts in CP, Id-1 and Id-2 immunoreactivity was as significantly elevated as in the cancer cells. These findings suggest that increased...

...Chemical Name: Binding Proteins; RNA, Messenger; Repressor Proteins; Transcription Factors; inhibitor of differentiation, helix-loop-helix protein; Id-2 protein; ID3 protein, human

9/3,K,AB/2 (Item 2 from file: 155)

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

11757557 99195162 PMID: 10095458

[Significance of differential nuclear expression of Ki-67 in adult soft tissue sarcomas]

Zur Bedeutung der differentiellen nuklearen Ki-67 Expression in Weichgewebssarkomen Erwachsener.

Rohr U P; Heinzinger M; Rheinlander B; Parwaresch R; Bohle R M  
Institut fur Pathologie, Universitat Giessen.

Verhandlungen der Deutschen Gesellschaft fur Pathologie (GERMANY)  
1998, 82 p345-50, ISSN 0070-4113 Journal Code: 7503704

Document type: Journal Article ; English Abstract

Languages: GERMAN

Main Citation Owner: NLM

Record type: Completed

As many other nuclear markers, e.g. steroid receptors, Ki-67 epitopes are differentially expressed in tumour cell nuclei. It is unclear whether this phenomenon represents tumour cell heterogeneity, different stages of the cell-cycle or a biological phenomenon with prognostic impact. We analysed 104 primary adult soft tissue sarcomas (ASTS), formalin-fixed, paraffin-embedded, by APAAP and LSAB immunohistochemistry, epitope retrieval technique and 2 anti-Ki-67 antibodies (MIB-1 and Ki-S-5). Expression was evaluated by 4 indexes/1000 tumour cells: a) A-index: sum of all (weak, moderate and strong stained) Ki-67-nuclei, b) the weighed R-index: sum of all strong stained Ki-67+ nuclei x3, moderate stained nuclei x2 and weak stained nuclei x1, c) ID1-index: sum of all strong stained Ki-67+ nuclei, and d) ID2 -index: sum of all strong and moderate stained Ki-67+ nuclei. Prognostic impact was analysed by Kaplan-Meier and logrank statistics with respect to overall survival. Quantitative Ki-67 expression did not vary significantly if determined by MIB-1 or Ki-S-5. The A-index turned out to be the strongest prognostic parameter within the whole group of ASTS as well as within each single sarcoma type investigated. Significant ( $p < 0.05$ ) correlations between A-index and overall survival existed in LMS, LPS, MFH, SS, while a trend to significance ( $p = 0.06$ ) was observed in MPNST. Quantitative evaluation of all three differential expression levels is necessary to obtain the most

**THIS PAGE BLANK (USPTO)**

comprehensive prognostic informations of proliferation markers in ASTS.

1998,

... paraffin-embedded, by APAAP and LSAB immunohistochemistry, epitope retrieval technique and 2 anti-Ki-67 antibodies (MIB-1 and Ki-S-5). Expression was evaluated by 4 indexes/1000 tumour cells...

... stained Ki-67+ nuclei x3, moderate stained nuclei x2 and weak stained nuclei x1, c) ID1-index: sum of all strong stained Ki-67+ nuclei, and d) ID2 -index: sum of all strong and moderate stained Ki-67+ nuclei. Prognostic impact was analysed...

9/3,K,AB/3 (Item 3 from file: 155)  
DIALOG(R)File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

11670622 99105716 PMID: 9890710

Characterization of two monoclonal antibodies against the RON tyrosine kinase receptor.

Montero-Julian F A; Dauny I; Flavetta S; Ronsin C; Andre F; Xerri L; Wang M H; Marvaldi J; Breathnach R; Brailly H

Immunotech 130, Marseille, France.

Hybridoma (UNITED STATES) Dec 1998, 17 (6) p541-51, ISSN 0272-457X Journal Code: 8202424

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

RON is a receptor protein tyrosine kinase belonging to the hepatocyte growth factor (HGF) receptor family. Using Recepteur d'Origine Nantaïs (RON) transfected cell lines, Macrophage Stimulating Protein (MSP) was identified as the ligand of RON. RON is synthesized as a single chain precursor, which subsequently is cleaved to yield a disulfide-linked heterodimer, with a 40-kDa alpha chain and a 150-kDa beta chain. Activation of RON by MSP results in cell migration, shape change, and proliferation. The present work centers on the production and characterization of two monoclonal antibodies (MAbs) to RON called ID-1 and ID-2. Antibodies were generated by immunization of mice with Madin-Darby Canine Kidney (MDCK) cells expressing human RON (clone RE7). Both antibodies recognized the mature and precursor form of RON. The specificity of the anti-RON antibodies was confirmed using a hepatocarcinoma cell line HepG2 expressing both task MET and RON receptors. Specific immunoprecipitation with ID-1 and ID-2 or anti-MET antibody followed by Western blotting under reducing conditions with rabbit polyclonal antibodies against RON and MET showed that our anti-RON antibodies recognize specifically the RON receptor. Ligand binding experiments showed that both antibodies are able to block the binding of radiolabeled MSP to RON and showed also that the antibodies recognize two different epitopes in the molecule. The blocking of MSP binding to RON by the anti-RON antibodies was confirmed by inhibition of cell migration induced by MSP in HT-29-D4 cells. Significant immunostaining was not observed in any subpopulation of whole blood with either ID-1 or ID-2. We analyzed the expression of RON receptor in a number of human hematopoietic and nonhematopoietic cells lines by flow cytometry. We found a strong mean of fluorescence intensity (MFI) in colon adenocarcinoma cells SW620 and HT-29-D4, low MFI in SVK14 and HepG2 cells, and no immunostaining in melanoma, lymphoma, and leukemia cells. Immunohistochemistry revealed that RON was expressed in germinal centers of tonsil, in skin, small intestine, and colon. These antibodies defined RON as CDw136 during the last leucocyte typing VI.

Characterization of two monoclonal antibodies against the RON

**THIS PAGE BLANK (USPTO)**

tyrosine kinase receptor.

Dec 1998,

... change, and proliferation. The present work centers on the production and characterization of two monoclonal antibodies (MAbs) to RON called ID-1 and ID-2. Antibodies were generated by immunization of mice with Madin-Darby Canine Kidney (MDCK) cells expressing human RON (clone RE7). Both antibodies recognized the mature and precursor form of RON. The specificity of the anti-RON antibodies was confirmed using a hepatocarcinoma cell line HepG2 expressing both task MET and RON receptors. Specific immunoprecipitation with ID-1 and ID-2 or anti-MET antibody followed by Western blotting under reducing conditions with rabbit polyclonal antibodies against RON and MET showed that our anti-RON antibodies recognize specifically the RON receptor. Ligand binding experiments showed that both antibodies are able to block the binding of radiolabeled MSP to RON and showed also that the antibodies recognize two different epitopes in the molecule. The blocking of MSP binding to RON by the anti-RON antibodies was confirmed by inhibition of cell migration induced by MSP in HT-29-D4 cells. Significant immunostaining was not observed in any subpopulation of whole blood with either ID-1 or ID-2. We analyzed the expression of RON receptor in a number of human hematopoietic and nonhematopoietic...

... RON was expressed in germinal centers of tonsil, in skin, small intestine, and colon. These antibodies defined RON as CDw136 during the last leucocyte typing VI.

Descriptors: Antibodies, Monoclonal--immunology--IM; \*Receptor Protein-Tyrosine Kinases--immunology--IM; \*Receptors, Cell Surface--immunology--IM; Antibodies, Monoclonal--analysis--AN; Antibody%% % Specificity; Cell Line; Dogs; Immunohistochemistry; Mice; Mice, Inbred BALB C; Rabbits; Radioligand Assay

Chemical Name: Antibodies, Monoclonal; Receptors, Cell Surface; RON protein; Receptor Protein-Tyrosine Kinases

9/3,K,AB/4 (Item 4 from file: 155)

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

11202145 98078766 PMID: 9418957

Helix-loop-helix proteins in Schwann cells: a study of regulation and subcellular localization of Ids, REB, and E12/47 during embryonic and postnatal development.

Stewart H J; Zoidl G; Rossner M; Brennan A; Zoidl C; Nave K A; Mirsky R; Jessen K R

Department of Anatomy, University College London, United Kingdom. ucgahes@ucl.ac.uk

Journal of neuroscience research (UNITED STATES) Dec 1 1997, 50

(5) p684-701, ISSN 0360-4012 Journal Code: 7600111

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Although basic helix-loop-helix (bHLH) proteins play an important role in transcriptional control in many cell types, the role of HLH proteins in Schwann cells has yet to be assessed. In this study, we have analyzed the expression of the dominant negative HLH genes, Id1 to Id4 and the class A gene REB, during Schwann cell development. We found that mRNA derived from these genes was present in the Schwann cell lineage throughout development including embryonic precursors and mature cells. The mRNA levels were not significantly regulated during development. Nevertheless, by using antibodies against the four different Id proteins, we found clear regulation of some of these genes at the protein level, in particular Id 2, 4, and REB, both in amount and nuclear/cytoplasmic

**THIS PAGE BLANK (USPTO)**

localization. All these proteins are found in the nuclei of Schwann cell precursors but are not seen in nuclei of Schwann cells of newborn nerves. We observed extensive overlap in Id expression, especially in Schwann cell precursors that co-expressed all four Id proteins and REB. We also showed that Id 1 and 2 were up-regulated as Schwann cells progressed through the cell cycle. These data indicate that HLH transcription factors act as regulators of Schwann cell development and point to the existence of as yet unidentified cell type-specific bHLH proteins in these cells.

Dec 1 1997,

... assessed. In this study, we have analyzed the expression of the dominant negative HLH genes, Id1 to Id4 and the class A gene REB, during Schwann cell development. We found that...

... and mature cells. The mRNA levels were not significantly regulated during development. Nevertheless, by using **antibodies** against the four different Id proteins, we found clear regulation of some of these genes at the protein level, in particular Id 2, 4, and REB, both in amount and nuclear/cytoplasmic localization. All these proteins are found...

... cell precursors that co-expressed all four Id proteins and REB. We also showed that Id 1 and 2 were up-regulated as Schwann cells progressed through the cell cycle. These data...

9/3,K,AB/5 (Item 5 from file: 155)  
DIALOG(R) File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

10514155 96324955 PMID: 8702531  
mRNA profiling of rat islet tumors reveals nkx 6.1 as a beta-cell-specific homeodomain transcription factor.

Jensen J; Serup P; Karlsen C; Nielsen T F; Madsen O D  
Hagedorn Research Institute, Niels Steensensvej 6, DK-2820 Gentofte, Denmark.

Journal of biological chemistry (UNITED STATES) Aug 2 1996, 271

(31) p18749-58, ISSN 0021-9258 Journal Code: 2985121R

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Development of a high capacity multiplex reverse transcriptase-polymerase chain reaction protocol has allowed us to screen lineage related rat islet tumors classified as alpha-, beta-, and delta-like as judged by their hormone profile for differential expression of more than 50 selected genes. We find that in addition to insulin the insulinoma express the normal beta-cell markers Pdx-1, IAPP, and Glut-2, and that these markers are absent from the glucagonoma: a reflection of the normal alpha-cell. Furthermore, this study suggests that the GLP-1, glucagon, GIP, IGF-1, and insulin receptors as well as E-cadherin, R-cadherin, Id-1, and Id-2 are differentially expressed within the islet of Langerhans. Importantly, insulinoma-specific expression of the recently cloned homeodomain protein Nkx 6.1 predicted beta-cell-specific expression in the normal islet. Immunohistochemistry using **antibodies** raised against recombinant Nkx 6.1 did indeed localize Nkx 6.1 expression exclusively to the nuclei of normal islet beta-cells. Apart from pancreatic islets only the antral part of the stomach contained Nkx 6.1 mRNA. We conclude that multiplex reverse transcriptase-polymerase chain reaction-based mRNA profiling is a powerful tool to identify differentially expressed genes within phenotypically related cells and propose that Nkx 6.1 is involved in specifying the unique characteristics of the beta-cell.

Aug 2 1996,

**THIS PAGE BLANK (USPTO)**

... 1, glucagon, GIP, IGF-1, and insulin receptors as well as E-cadherin, R-cadherin, Id-1, and Id-2 are differentially expressed within the islet of Langerhans. Importantly, insulinoma-specific expression of the recently...

... protein Nkx 6.1 predicted beta-cell-specific expression in the normal islet. Immunohistochemistry using **antibodies** raised against recombinant Nkx 6.1 did indeed localize Nkx 6.1 expression exclusively to ...

9/3,K,AB/6 (Item 6 from file: 155)  
DIALOG(R) File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

08261177 94327199 PMID: 7519580

Monoclonal **antibody** against the active site of caeruloplasmin and the ELISA system detecting active caeruloplasmin.

Hiyamuta S; Ito K

Central Research Laboratories Idemitsu Kosan Co., Ltd., Chiba, Japan.

Hybridoma (UNITED STATES) Apr 1994, 13 (2) p139-41, ISSN

0272-457X Journal Code: 8202424

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Serum caeruloplasmin deficiency is a characteristic biochemical abnormality found in patients with Wilson's disease, but the mechanism of this disease is unknown. Although the phenylenediamine oxidase activity of serum caeruloplasmin is markedly low in patients with Wilson's disease, mRNA of caeruloplasmin exists to some extent. To investigate the deficiency of caeruloplasmin oxidase activity in Wilson's disease, we generated 14 monoclonal **antibodies** (MAbs) and selected ID1, which had the strongest reactivity, and ID2, which had neutralizing ability. We also established a system to measure active caeruloplasmin specifically using these MAbs. These MAbs and the system will be useful tools in analyzing the active site of caeruloplasmin in patients with Wilson's disease.

Monoclonal **antibody** against the active site of caeruloplasmin and the ELISA system detecting active caeruloplasmin.

Apr 1994,

... investigate the deficiency of caeruloplasmin oxidase activity in Wilson's disease, we generated 14 monoclonal **antibodies** (MAbs) and selected ID1, which had the strongest reactivity, and ID2, which had neutralizing ability. We also established a system to measure active caeruloplasmin specifically using...

Descriptors: **Antibodies**, Monoclonal--immunology--IM; \*Ceruloplasmin --immunology--IM; \*Epitopes--immunology--IM

Chemical Name: **Antibodies**, Monoclonal; Epitopes; Ceruloplasmin

9/3,K,AB/7 (Item 7 from file: 155)  
DIALOG(R) File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

08041575 94107309 PMID: 8280128

Lack of copper binding sites in ceruloplasmin of LEC rats with abnormal copper metabolism.

Hiyamuta S; Takeichi N

Central Research Laboratories, Idemitsu Kosan Co., Ltd., Chiba, Japan.

Biochemical and biophysical research communications (UNITED STATES) Dec 30 1993, 197 (3) p1140-5, ISSN 0006-291X Journal Code: 0372516

Document type: Journal Article

**THIS PAGE BLANK (USPTO,**

Languages: ENGLISH  
Main Citation Owner: NLM  
Record type: Completed

Recently it was found that the clinical features of the LEC rat closely resemble those of human Wilson's disease. One of the characteristics of the animal is low levels of serum ceruloplasmin. Therefore, by using LEC rats, we attempted to define molecular basis of the deficiency in active site of ceruloplasmin in Wilson's disease patients. We made 3 monoclonal **antibodies**, ID2 against active site of ceruloplasmin, ID1 against inactive site of ceruloplasmin, and the remaining one against metallothionein. Using these monoclonal **antibodies**, we examined immunohistochemical stainings of LEC rat liver tissues, and compared them with those of LEA rats, as a control. ID1 stained the hepatocytes of both LEA and LEC rats, whereas ID2 stained LEA rat hepatocytes only. The results indicated that the ceruloplasmin secreted by LEC rat hepatocytes is mostly in inactive form. The **antibody** against metallothionein stained LEA rat hepatocytes only. This finding may also indicate that LEC rat hepatocytes express less amount of metallothionein than those of LEA rats.

Dec 30 1993,

... deficiency in active site of ceruloplasmin in Wilson's disease patients. We made 3 monoclonal **antibodies**, ID2 against active site of ceruloplasmin, ID1 against inactive site of ceruloplasmin, and the remaining one against metallothionein. Using these monoclonal **antibodies**, we examined immunohistochemical stainings of LEC rat liver tissues, and compared them with those of LEA rats, as a control. ID1 stained the hepatocytes of both LEA and LEC rats, whereas ID2 stained LEA rat hepatocytes only. The results indicated that the ceruloplasmin secreted by LEC rat hepatocytes is mostly in inactive form. The **antibody** against metallothionein stained LEA rat hepatocytes only. This finding may also indicate that LEC rat...

; **Antibodies**, Monoclonal; Binding Sites; Ceruloplasmin--chemistry --CH; Hepatolenticular Degeneration--metabolism--ME; Immunohistochemistry; Liver--pathology--PA; Metal...

Chemical Name: **Antibodies**, Monoclonal; Copper; Metallothionein; Ceruloplasmin

9/3,K,AB/8 (Item 8 from file: 155)  
DIALOG(R) File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

06396719 90021152 PMID: 2508304

Interdigitating cell sarcoma (ICS). Evidence of interdigitating cell origin, immunocytochemical studies with monoclonal anti-ICS **antibodies**.

Nakamura S; Suchi T; Suzuki R; Takagi N; Kitoh K; Osada H; Ueda R; Takahashi T; Hiai H; Kato K; et al

Department of Pathology, Aichi Cancer Center Hospital, Nagoya, Japan.

Virchows Archiv. A, Pathological anatomy and histopathology (GERMANY, WEST) 1989, 415 (5) p447-57, ISSN 0174-7398 Journal Code: 8302198

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Three independent mouse monoclonal **antibodies** (mAbs) ID1 (IgG3), ID2 and ID3 (IgM) were raised against whole cells of a surgically resected human interdigitating cell sarcoma (ICS). In immunoperoxidase staining, these mAbs strongly stained the cytoplasm of ICS neoplastic cells as well as interdigitating cells in normal lymphoid tissues. These mAbs also detected monocyte/macrophages and dendritic cells, although their staining was highly variable depending on tissue

**THIS PAGE BLANK (USPTO)**

distribution of the cells. Additional immuno-histological and enzyme histochemical study revealed that the neoplastic cells of ICS had cytoplasmic acid phosphatase and membranous alkaline phosphatase activity, and also possessed S100 beta protein, Ki-1 antigen. DAKO-macrophage antigen, and weak vimentin activity. Neither rearrangement of immunoglobulin heavy chain gene nor of T-cell receptor genes was detected in the DNA of ICS by Southern hybridization. These observations provide further confirmation of our previous finding (Nakamura et al. 1988, 1989) that the origin of ICS is interdigitating rather than lymphoid cell, and indicate that our mAbs could be useful as a cellular differentiation marker of interdigitating cells and for diagnosis of ICS.

Interdigitating cell sarcoma (ICS). Evidence of interdigitating cell origin, immunocytochemical studies with monoclonal anti-ICS antibodies.

1989,

Three independent mouse monoclonal antibodies (mAbs) ID1 (IgG3), ID2 and ID3 (IgM) were raised against whole cells of a surgically resected human interdigitating cell...

Descriptors: **Antibodies**, Monoclonal--diagnostic use--DU; \*Antigens, Neoplasm--analysis--AN; \*Sarcoma--pathology--PA

Chemical Name: **Antibodies**, Monoclonal; Antigens, Neoplasm

9/3,K,AB/9 (Item 9 from file: 155)

DIALOG(R) File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

05572996 87252223 PMID: 3110274

Molecular analysis of heavy and light chains used by primary and secondary anti-(T,G)-A--L **antibodies** produced by normal and xid mice.

Busto P; Gerstein R; Dupre L; Giorgetti C A; Selsing E; Press J L

Journal of immunology (Baltimore, Md. - 1950) (UNITED STATES) Jul 15 1987, 139 (2) p608-18, ISSN 0022-1767 Journal Code: 2985117R

Contract/Grant No.: AI-13725; AI; NIAID

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

The primary (1 degree) **antibody** response to (T,G)-A--L shows limited heterogeneity, consisting mostly of side chain-specific **antibodies** that bind GT and that express the TGB5 idiotype (Id). The secondary (2 degrees) response is very diverse: **antibodies** that bind the backbone A--L constitute a third of the response, and a high proportion of the side chain-specific **antibodies** do not bind GT and are TGB5 Id-. To provide a molecular basis for understanding this difference in repertoire expression, we analyzed the Ig genes used by heavy and light chains of 1 degree and 2 degrees side chain-specific anti-(T,G)-A--L hybridoma **antibodies** (HP). Southern blot restriction analysis and nucleotide sequence analysis of the expressed genes used by three TGB5 Id+ 2 degrees HP showed usage of three different VH genes in two VH gene families (36-60 and J558), different D segments, and two different Vk1 genes (the Vk1A and Vk1C subgroups). Thus, **antibody** heterogeneity in the 2 degrees response is contributed by combinatorial diversity of distinct germ-line genes. Nucleotide sequence analysis of the expressed genes used by TGB5 Id+ 1 degree HP showed use of highly homologous VH genes in the J558 VH gene family and highly homologous Vk1A genes. The majority of TGB5 Id+ 1 degree HP from different donors gave similar heavy and similar light chain gene rearrangements by Southern blot restriction analysis, after correction for known or potential J region differences. The combined nucleotide sequence and Southern blot restriction analysis data suggest that most 1 degree B cells use the same or very similar VH and Vk genes, i.e., the 1 degree response is paucigenic. Different D segments were used by the TGB5 Id+ 1 degree and 2

**THIS PAGE BLANK (USPTO)**

degrees HP that were sequenced, and there was no apparent correlation between TGB5 idiotype and VH, D gene, or JH gene usage. However, all TGB5 Id+ HP sequenced used highly homologous genes from the Vk1 group. Expression of a Vk1 light chain correlates with, but is not sufficient for, TGB5 idiotype, because one GT-binding, TGB5 Id- HP was found to use a Vk1C subgroup light chain. By Southern blot and nucleotide sequence analysis, the Vk genes used by two TGB5 Id+ 2 degrees HP from xid mice are highly homologous, if not identical to the Vk1A gene(s) used by 1 degree and 2 degrees Id+ HP from wild-type mice.

... of heavy and light chains used by primary and secondary anti-(T,G)-A--L antibodies produced by normal and xid mice.

Jul 15 1987,

The primary (1 degree) antibody response to (T,G)-A--L shows limited heterogeneity, consisting mostly of side chain-specific antibodies that bind GT and that express the TGB5 idiotype (Id). The secondary (2 degrees) response is very diverse: antibodies that bind the backbone A--L constitute a third of the response, and a high proportion of the side chain-specific antibodies do not bind GT and are TGB5 Id-. To provide a molecular basis for understanding...

...of 1 degree and 2 degrees side chain-specific anti-(T,G)-A--L hybridoma antibodies (HP). Southern blot restriction analysis and nucleotide sequence analysis of the expressed genes used by three TGB5 Id+ 2 degrees HP showed usage of three different VH genes in two VH gene families (36...

... J558), different D segments, and two different Vk1 genes (the Vk1A and Vk1C subgroups). Thus, antibody heterogeneity in the 2 degrees response is contributed by combinatorial diversity of distinct germ-line genes. Nucleotide sequence analysis of the expressed genes used by TGB5 Id+ 1 degree HP showed use of highly homologous VH genes in the J558 VH gene family and highly homologous Vk1A genes. The majority of TGB5 Id+ 1 degree HP from different donors gave similar heavy and similar light chain gene rearrangements by...

... e., the 1 degree response is paucigenic. Different D segments were used by the TGB5 Id+ 1 degree and 2 degrees HP that were sequenced, and there was no apparent correlation between...

... chain. By Southern blot and nucleotide sequence analysis, the Vk genes used by two TGB5 Id+ 2 degrees HP from xid mice are highly homologous, if not identical to the Vk1A gene...

9/3,K,AB/10 (Item 10 from file: 155)  
DIALOG(R)File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Corp. All rts. reserv.

05417511 87096023 PMID: 3541426

Intradermal hepatitis B vaccination in an abbreviated schedule.

Halsey N A; Reppert E J; Margolis H S; Francis D P; Fields H A

Vaccine (ENGLAND) Dec 1986, 4 (4) p228-32, ISSN 0264-410X

Journal Code: 8406899

Document type: Clinical Trial; Controlled Clinical Trial; Journal Article  
; Randomized Controlled Trial

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Two low-dose intradermal regimens for hepatitis B vaccination were compared with the standard 1 ml dose administered intramuscularly to healthy, 22-42 year old individuals. All regimens were administered in an abbreviated time schedule. Nineteen individuals (ID-1 group) received three 0.1 ml (2 micrograms) doses intradermally at times 0, 1

**THIS PAGE BLANK (USPTO)**

month and 4 months. Twenty-four individuals (ID-2 group) received two injections of 0.2 ml (4 micrograms) each intradermally at time 0 and one 0.1 ml (2 micrograms) injection 4 months later. Twenty individuals (IM group) received the recommended three 1.0 ml (20 micrograms) doses intramuscularly at times 0, 1 month, and 4 months. No significant adverse reactions were attributable to the intradermal administration of vaccine although the majority of vaccinees developed small areas of induration and hyperpigmentation at the injection site that persisted for several months. One month following the last injection, all vaccinees had developed anti-HBsAg **antibodies**. One hundred percent of ID-1 and IM vaccinees and 95% of ID-2 vaccinees had protective levels of **antibody** (greater than or equal to 10 mIU ml<sup>-1</sup>). The geometric mean titre (GMT) for the IM group (2692 mIU ml<sup>-1</sup>) was somewhat higher than for the ID-1 (1230 mIU ml<sup>-1</sup>) and the ID-2 (851 mIU ml<sup>-1</sup>) groups, but the differences were not statistically significant. Since anti-HBs **antibodies** are thought to confer protection against hepatitis B, these results suggest that a shortened regimen of intradermal vaccine may be effective in healthy adults. However, no efficacy study has yet been done with intradermal hepatitis B vaccine.

Dec 1986,

... 42 year old individuals. All regimens were administered in an abbreviated time schedule. Nineteen individuals (ID-1 group) received three 0.1 ml (2 micrograms) doses intradermally at times 0, 1 month and 4 months. Twenty-four individuals (ID-2 group) received two injections of 0.2 ml (4 micrograms) each intradermally at time 0...

... for several months. One month following the last injection, all vaccinees had developed anti-HBsAg **antibodies**. One hundred percent of ID-1 and IM vaccinees and 95% of ID-2 vaccinees had protective levels of **antibody** (greater than or equal to 10 mIU ml<sup>-1</sup>). The geometric mean titre (GMT) for the IM group (2692 mIU ml<sup>-1</sup>) was somewhat higher than for the ID-1 (1230 mIU ml<sup>-1</sup>) and the ID-2 (851 mIU ml<sup>-1</sup>) groups, but the differences were not statistically significant. Since anti-HBs **antibodies** are thought to confer protection against hepatitis B, these results suggest that a shortened regimen...

; Adult; Clinical Trials; Erythema--etiology--ET; Hepatitis B **Antibodies**--biosynthesis--BI; Injections, Intradermal; Random Allocation; Viral Hepatitis Vaccines--adverse effects--AE

Chemical Name: Hepatitis B **Antibodies**; Viral Hepatitis Vaccines

9/3,K,AB/11 (Item 1 from file: 34)

DIALOG(R)File 34:SciSearch(R) Cited Ref Sci

(c) 2003 Inst for Sci Info. All rts. reserv.

04579028 Genuine Article#: TU871 Number of References: 37

Title: MULTIPLE DOMAINS CONTRIBUTE TO THE DISTINCT INACTIVATION PROPERTIES OF HUMAN HEART AND SKELETAL-MUSCLE NA<sup>+</sup> CHANNELS (Abstract Available)

Author(s): MAKITA N; BENNETT PB; GEORGE AL

Corporate Source: VANDERBILT UNIV,MED CTR,S-3223 MCN,21ST AVE S &GARLAND AVE/NASHVILLE//TN/37232; VANDERBILT UNIV,SCH MED,DEPT MED/NASHVILLE//TN/37212; VANDERBILT UNIV,SCH MED,DEPT PHARMACOL/NASHVILLE//TN/37212

Journal: CIRCULATION RESEARCH, 1996, V78, N2 (FEB), P244-252

ISSN: 0009-7330

Language: ENGLISH Document Type: ARTICLE

Abstract: Voltage-gated Na<sup>+</sup> channels are essential for the normal electrical excitability of neuronal and striated muscle membranes. Distinct isoforms of the Na<sup>+</sup> channel alpha-subunit have been identified by molecular cloning, and their functional attributes have been defined by heterologous expression coupled with electrophysiological

**THIS PAGE BLANK (USPTO)**

recording. Two closely related Na<sup>+</sup> channel alpha-subunit isoforms, hH1 (human heart) and hSkM1 (human skeletal muscle), exhibit differences in their inactivation properties and in their response to the coexpressed beta(1)-subunit. To localize regions that contribute to inactivation and to beta(1)-subunit response, we have exploited these functional differences by studying chimeric channels composed of segments from both hH1 and hSkM1. Chimeras in which one or more of the cytoplasmic interdomain regions (ID1-2, ID2-3, and ID3-4) were exchanged between hH1 and hSkM1 exhibit inactivation properties identical with the background channel isoform, suggesting that these regions are not sufficient to cause gating differences. In contrast, inactivation properties of chimeras composed of approximately equal halves of the two channel isoforms were intermediate between hH1 and hSkM1. Furthermore, the response to the coexpressed beta(1)-subunit was dependent on structures located in the carboxy-terminal half of the ac-subunit, although domains D3, D4, and the carboxy terminal are not singularly responsible for this effect. These data indicate that inactivation differences between hH1 and hSkM1 are determined by multiple alpha-subunit domains.

, 1996

...Abstract: both hH1 and hSkM1. Chimeras in which one or more of the cytoplasmic interdomain regions (ID1-2, ID2-3, and ID3-4) were exchanged between hH1 and hSkM1 exhibit inactivation properties identical with...

...Identifiers--DEPENDENT SODIUM-CHANNEL; FUNCTIONAL EXPRESSION; PERIODIC PARALYSIS; RAT SKELETAL; **ANTIBODIES**; RECEPTOR; SUBUNITS; BETA-1; SITE

9/3,K,AB/12 (Item 1 from file: 340)  
DIALOG(R) File 340:CLAIMS(R)/US Patent  
(c) 2003 IFI/CLAIMS(R). All rts. reserv.

Dialog Acc No: 3093770 IFI Acc No: 9900471

Document Type: C

METHOD FOR DETECTING IMMUNE RESPONSE TO HEPATITIS B; USING AN OLIGO(OR POLY) PEPTIDE

Inventors: Thakur Arvind (US); Thanavala Yasmin (US)

Assignee: Health Research Inc; London, University College GB

Assignee Code: 11003 11684

Publication (No,Date), Applic (No,Date):

US 5856087 19990105 US 97948762 19971010

Publication Kind: A

Calculated Expiration: 20131215

Continuation Pub(No), Applic(No,Date): US 5531990 US 93167336

19931215; US 5744135 US 96589011 19960119

Priority Applic(No,Date): US 97948762 19971010; US 93167336 19931215;

US 96589011 19960119

Abstract: The invention comprises an anti-idiotypic **antibody** designated 2F10 and permitted variants thereof, which have antigenic properties similar to the group specific 'a' determinant of human hepatitis B surface antigen HBsAg and have at least partial but not complete homology with such surface antigen. The invention further comprises a peptide having a chain comprising the amino acid residues Ala Val Tyr Tyr Cys Thr Arg Gly Tyr His Gly Ser Ser Leu Tyr and permitted variants thereof, which, like 2F10, have antigenic properties similar to the group specific 'a' determinant of human hepatitis B surface antigen HBsAg and have at least partial, but not complete, homology with said surface antigen. The amino acid sequence is found in and forms a part of 2F10. The shorter peptide chain comprising the amino acid residues Gly Tyr His Gly Ser Ser Leu Tyr and permitted variants thereof, also have antigenic properties similar to the group specific 'a' determinant of human

**THIS PAGE BLANK (USPTO)**

hepatitis B surface antigen HBsAg and have at least partial, but not complete, homology with said surface antigen.

Publication (No,Date), Applic (No,Date):  
...19990105

Abstract: The invention comprises an anti-idiotypic **antibody** designated 2F10 and permitted variants thereof, which have antigenic properties similar to the group specific...

Exemplary Claim: ...Val Tyr Tyr Cys Thr Arg Gly Tyr His Gly Ser Ser Leu Tyr (Sequence ID #1) and detecting a response to said sequence.

Non-exemplary Claims: ...with the 8 amino acid sequence Gly Tyr His Gly Ser Ser Leu Tyr (Sequence ID #2) and detecting a response to said sequence.

9/3,K,AB/13 (Item 2 from file: 340)  
DIALOG(R) File 340:CLAIMS(R)/US Patent  
(c) 2003 IFI/CLAIMS(R). All rts. reserv.

Dialog Acc No: 2930345 IFI Acc No: 9801537

Document Type: C

MONONUCLEAR LEUKOCYTE DIRECTED ENDOTHELIAL ADHESION MOLECULE ASSOCIATED WITH ATHEROSCLEROSIS

Inventors: Collins Tucker (US); Cybulsky Myron I (US); Gimbrone Michael A Jr (US)

Assignee: Brigham and Women's Hospital

Assignee Code: 08822

Publication (No,Date), Applic (No,Date):

US 5708147 19980113 US 94261304 19940616

Publication Kind: A

Calculated Expiration: 20150113

Document Type: CERTIFICATE OF CORRECTION Certificate of Correction Date:  
19980818

Continuation Pub(No), Applic(No,Date): ABANDONED US 91649565  
19910201

Cont.-in-part Pub(No), Applic(No,Date): ABANDONED US  
90487038 19900302

Priority Applic(No,Date): US 94261304 19940616; US 91649565 19910201;  
US 90487038 19900302

Abstract: The invention relates to novel endothelial cell-leukocyte adhesion molecules designated ATHERO-ELAM. ATHERO-ELAM molecules are expressed on cultured endothelial cells stimulated with bacterial LPS and selectively mediate the binding of monocytes to the endothelial cells. Monoclonal **antibodies** specific for ATHEROELAM bind to vascular endothelial cells involved in early atherosclerotic lesions, but not to vascular endothelial cells from uninvolved arterial tissue. ATHERO-ELAM and **antibodies** directed to ATHERO-ELAM may be used in identifying early atherosclerotic lesions and in treating and preventing atherosclerosis.

Publication (No,Date), Applic (No,Date):  
...19980113

Abstract: ...with bacterial LPS and selectively mediate the binding of monocytes to the endothelial cells. Monoclonal **antibodies** specific for ATHEROELAM bind to vascular endothelial cells involved in early atherosclerotic lesions, but not to vascular endothelial cells from uninvolved arterial tissue. ATHERO-ELAM and **antibodies** directed to ATHERO-ELAM may be used in identifying early atherosclerotic lesions and in treating...

**THIS PAGE BLANK (USPTO)**

Exemplary Claim: ...leukocyte adhesion molecule expressed in atherosclerotic lesions having the amino acid sequence shown in Sequence ID 2, and comprising an AS-1 domain between domains 3 and 4 of said protein and...

Non-exemplary Claims: ...claim 1 wherein said protein is encoded by the nucleotide sequence as shown in Sequence ID 1.

...

...domains, wherein said protein comprises the sequence between amino acids 1 and 774 in Sequence ID 2, and wherein said protein is selected from the group consisting of: a protein having seven

9/3,K,AB/14 (Item 3 from file: 340)  
DIALOG(R)File 340:CLAIMS(R)/US Patent  
(c) 2003 IFI/CLAIMS(R). All rts. reserv.

Dialog Acc No: 2884574 IFI Acc No: 9726126

Document Type: C

DETERGENT COMPOUNDS WITH HIGH ACTIVITY CELLULASE AND QUATERNARY AMMONIUM COMPOUNDS; CATIONIC SURFACTANTS AND CELLULASES FOR LAUNDRY DETERGENTS

Inventors: Baeck Andre Cesar (BE); Busch Alfred (BE); Convents Andre Christian (BE)

Assignee: Procter & Gamble Co The

Assignee Code: 68128

Publication (No,Date), Applic (No,Date):

US 5668073 19970916 US 96666147 19960619

Publication Kind: A

Calculated Expiration: 20141117

(Cited in 001 later patents)

Continuation Pub(No),Applic(No,Date):

19941117

US 94290712

Priority Applic(No,Date): EP 91202881 19911106

Abstract: The present invention provides a detergent composition comprising a quaternary ammonium compound of formula:  $R_1R_2R_3R_4N^+X^-$ , wherein  $R_1$  is C8-C16 alkyl, each of  $R_2$ ,  $R_3$  and  $R_4$  is independently C1-C4 alkyl or hydroxy alkyl, benzyl or  $-(C_2H_4O)_xH$  where  $x$  has a value from 2 to 5, not more of  $R_2$ ,  $R_3$  or  $R_4$  being benzyl, and  $X$  is an anion, and a cellulase characterized in that said cellulase provides at least 10% removal of immobilized radio-active labelled carboxymethylcellulose according to the CMC-method at  $25 \times 10^{-6}\%$  by weight of cellulase protein in the laundry test solution. According to the present invention, a preferred cellulase consists of a homogeneous endoglucanase component which is immunoreactive with a monoclonal **antibody** raised against a partially purified 43 kD cellulase derived from *Humicola insolens* DM 1800.

Publication (No,Date), Applic (No,Date):

...19970916

Abstract: ...a preferred cellulase consists of a homogeneous endoglucanase component which is immunoreactive with a monoclonal **antibody** raised against a partially purified 43 kD cellulase derived from *Humicola insolens* DM 1800.

Exemplary Claim: ...wherein said cellulase consists essentially of a homogeneous endoglucanase component which is immunoreactive with an **antibody** raised against a highly purified about 43 kD cellulase derived from *Humicola insolens*, DSM 1800.

Non-exemplary Claims: ...1 wherein the cellulase has the amino acid sequence shown in the appended sequence listing ID #2, or is a homologue thereof exhibiting endoglucanase activity...

**THIS PAGE BLANK (USPTO)**

...to claim 8 wherein the DNA sequence is as shown in the appended sequence listings ID #1 or ID #3...

9/3,K,AB/15 (Item 4 from file: 340)  
DIALOG(R)File 340:CLAIMS(R)/US Patent  
(c) 2003 IFI/CLAIMS(R). All rts. reserv.

Dialog Acc No: 2722816 IFI Acc No: 9612741  
Document Type: C  
COMPACT DETERGENT COMPOSITIONS WITH HIGH ACTIVITY CELLULASE; SURFACTANTS,  
BUILDERS AND CELLULASE IN PARTICLES  
Inventors: Baeck Andre C (BE); Busch Alfred (BE); Ceulemans Raphael A (BE)  
Assignee: Procter & Gamble Co The  
Assignee Code: 68128  
Publication (No,Date), Applic (No,Date):  
US 5520838 19960528 US 9381328 19931119  
Publication Kind: A  
Calculated Expiration: 20130528  
(Cited in 004 later patents) Document Type: CERTIFICATE OF CORRECTION  
Certificate of Correction Date: 19961203  
PCT Pub(No,Date), Applic(No,Date): WO 915841 19910205 WO  
92US203 19920115  
Section 371: 19931119  
Section 102(e):19931119  
Priority Applic(No,Date): EP 91870006 19910116; EP 91202879 19911106

Abstract: The present invention concerns cellulase-containing granular detergent compositions which are in a 'compact' form, i.e. they are of a relatively high density and contain a relatively low amount of inorganic filler salt compared to conventional detergent compositions. In the detergent compositions herein the cellulase is defined by the C14CMC method described herein and preferably comprises a specific single-component endoglucanase.

Publication (No,Date), Applic (No,Date):  
...19960528  
...PCT Pub(No,Date), Applic(No,Date): 19910205

Exemplary Claim: ...said cellulase consists essentially of a homogeneous endoglucanase component which is immunoreactive with a monoclonal antibody raised against a partially purified about 43 kD cellulase derived from Humicola insolens, DSM 1800...

Non-exemplary Claims: ...is an endoglucanase enzyme having the amino acid sequence shown in the appended sequence listing ID#2; said granular detergent composition comprising no more than about 15% by weight of inorganic filler...

...to claim 15 wherein the DNA sequence is as shown in the appended sequence listings ID #1 or ID #3...

?  
PLEASE ENTER A COMMAND OR BE LOGGED OFF IN 5 MINUTES  
?

**THIS PAGE BLANK (USPTO)**

278752 21015716 PMID: 11131972

PROTEAN. Protein sequence analysis and prediction.

Plasterer T N

Biomolecular Engineering Resource Center, Boston University, Boston, MA 02215, USA. tplas@bu.edu

Molecular biotechnology (United States) Oct 2000, 16 (2) p117-25, ISSN 1073-6085 Journal Code: 9423533

Document type: Journal Article; Review; Review, Tutorial

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

The archaeal, bacterial, and eukaryotic genome projects have overwhelmed our ability to experimentally elucidate the **function** of each novel gene and gene product. To a certain extent, protein functional assignments can be derived via sequence similarity measures and direct primary sequence analysis using methods to predict hydropathy, secondary structure, amphiplicity, and antigenicity. **Function** can also be inferred on the basis of sequence motifs, such as phosphorylation and lipid binding signatures. These methods, provided in DNASTAR's PROTEAN module, can be used to putatively assign roles for novel proteins from the genome explosion as well as clarify **function** for better known proteins.

The archaeal, bacterial, and eukaryotic genome projects have overwhelmed our ability to experimentally elucidate the **function** of each novel gene and gene product. To a certain extent, protein functional assignments can...

**THIS PAGE BLANK (USPTO)**

SYSTEM:OS - DIALOG OneSearch

File 155:MEDLINE(R) 1966-2003/Oct W4

(c) format only 2003 The Dialog Corp.

\*File 155: Please see HELP NEWS 155 for details about the 2003 reload.

File 55:BIOSIS Previews(R) 2003-2003/Oct W4

(c) 2003 BIOSIS

\*File 55: BIOSIS Previews has been reloaded with major enhancements.

See HELP NEWS055 for more information.

File 34:SciSearch(R) Cited Ref Sci 1990-2003/Oct W3

(c) 2003 Inst for Sci Info

File 434:SciSearch(R) Cited Ref Sci 1974-1989/Dec

(c) 1998 Inst for Sci Info

File 340:CLAIMS(R)/US Patent 1950-03/Oct 28

(c) 2003 IFI/CLAIMS(R)

\*File 340: Enter HELP NEWS340 & HELP ALERTS340 for search,  
display & Alert information.

Set Items Description

--- ----

? s dnastar

S1 74 DNASTAR

? s activity or function

Processing

3149008 ACTIVITY

2363016 FUNCTION

S2 5171310 ACTIVITY OR FUNCTION

? s s1 and s2

74 S1

5171310 S2

S3 7 S1 AND S2

? rd

>>>Duplicate detection is not supported for File 340.

>>>Records from unsupported files will be retained in the RD set.

...completed examining records

S4 5 RD (unique items)

? t s4/3,k,ab/1-5

4/3,K,AB/1 (Item 1 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

15076128 22561985 PMID: 12674638

Cloning and expression product of vip3A gene from Bacillus thuringiensis  
and analysis of insecticidal activity]

Chen Jian-Wu; Tang Li-Xia; Tang Mu-Jin; Shi Yong-Xia; Pang Yi

State Key Laboratory for Biocontrol, Zhongshan University, Guangzhou  
510275, China.

Sheng wu gong cheng xue bao = Chinese journal of biotechnology (China)

Nov 2002, 18 (6) p687-92, ISSN 1000-3061 Journal Code: 9426463

Document type: Journal Article ; English Abstract

Languages: CHINESE

Main Citation Owner: NLM

Record type: Completed

The vip3 A gene in a size of 2.3 kb amplified from wild-type Bacillus  
thuringiensis strain S184 by PCR was cloned into pGEM-T Easy vector and its  
sequence was analyzed by DNASTAR. The plasmid pOTP was constructed  
by inserting vip3A-S184 gene into the expression vector pQE30 and then was  
transformed into E. coli M15. E. coli M15 cells harbouring the plasmid pOTP  
were induced with 1 mmol/L IPTG to express 89 kD protein which was  
confirmed to be Vip3A-S184 by Western blot. Experiments showed that about  
19% of Vip3A-S184 proteins were soluble, and others were insoluble proteins  
and formed inclusion bodies observed by transmission electron  
microscopy(TEM). The target protein was purified under the native condition

**THIS PAGE BLANK (USPTO)**

and the polyclonal antibody was prepared by immunizing rabbits. The polyclonal antibody was used to detect Vip3A proteins expressed in *Bacillus thuringiensis*. Bioassay showed that Vip3A-S184 showed a high toxicity against 3 tested insect larvae including *Spodoptera exigua*, *Spodoptera litura* and *Helicoverpa armigera*.

Cloning and expression product of vip3A gene from *Bacillus thuringiensis* and analysis of insecticidal activity]

... by PCR was cloned into pGEM-T Easy vector and its sequence was analyzed by DNASTAR. The plasmid pOTP was constructed by inserting vip3A-S184 gene into the expression vector pQE30...

4/3,K,AB/2 (Item 2 from file: 155)  
DIALOG(R)File 155:MEDLINE(R)  
(c) format only 2003 The Dialog Cor

**THIS PAGE BLANK (USPTO)**

...METHODS & EQUIPMENT: DNASTar program

? ds

Set	Items	Description
S1	74	DNASTAR
S2	5171310	ACTIVITY OR FUNCTION
S3	7	S1 AND S2
S4	5	RD (unique items)

? s review

S5 881558 REVIEW

? s s1 and s5

74	S1
881558	S5
S6	6 S1 AND S5

? rd

>>>Duplicate detection is not supported for File 340.

>>>Records from unsupported files will be retained in the RD set.

...completed examining records

S7 5 RD (unique items)

? t s7/3,k,ab/1-5

7/3,K,AB/1 (Item 1 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

10230537 96031839 PMID: 7552691

Macintosh sequence analysis software. DNASTar's LaserGene.

Clewley J P

Virus Reference Division, Central Public Health Laboratory, London, UK.

Molecular biotechnology (UNITED STATES) Jun 1995, 3 (3) p221-4,

ISSN 1073-6085 Journal Code: 9423533

Document type: Journal Article; Review; Review, Tutorial

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

The analysis of information in nucleotide and amino acid sequence data from an investigator's own laboratory, or from the ever-growing worldwide databases, is critically dependent on well planned and written software. Although the most powerful packages previously have been confined to workstations, there has been a dramatic increase over the last few years in the sophistication of the programs available for personal computers, as the speed and power of these have increased. A wide choice of software is available for the Macintosh, including the LaserGene suite of programs from DNASTar. This review assesses the strengths and weaknesses of LaserGene and concludes that it provides a useful and comprehensive range of sequence analysis tools.

Macintosh sequence analysis software. DNASTar's LaserGene.

... choice of software is available for the Macintosh, including the LaserGene suite of programs from DNASTar. This review assesses the strengths and weaknesses of LaserGene and concludes that it provides a useful and...

7/3,K,AB/2 (Item 1 from file: 55)

DIALOG(R)File 55:Biosis Previews(R)

(c) 2003 BIOSIS. All rts. reserv.

0012883107 BIOSIS NO.: 200100054946

PROTEAN: Protein sequence analysis and prediction

AUTHOR: Plasterer Thomas N (Reprint)

AUTHOR ADDRESS: Biomolecular Engineering Resource Center, Boston

University, Boston, MA, 02215, USA\*\*USA

JOURNAL: Molecular Biotechnology 16 (2): p117-125 October, 2000 2000

**THIS PAGE BLANK (USPTO)**

MEDIUM: print  
ISSN: 1073-6085  
DOCUMENT TYPE: Article; Literature Review  
RECORD TYPE: Abstract  
LANGUAGE: English

ABSTRACT: The archaeal, bacterial, and eukaryotic genome projects have overwhelmed our ability to experimentally elucidate the function of each novel gene and gene product. To a certain extent, protein functional assignments can be derived via sequence similarity measures and direct primary sequence analysis using methods to predict hydropathy, secondary structure, amphiphilicity, and antigenicity. Function can also be inferred on the basis of sequence motifs, such as phosphorylation and lipid binding signatures. These methods, provided in DNASTAR's PROTEAN module, can be used to putatively assign roles for novel proteins from the genome explosion as well as clarify function for better known proteins.

...ABSTRACT: basis of sequence motifs, such as phosphorylation and lipid binding signatures. These methods, provided in DNASTAR's PROTEAN module, can be used to putatively assign roles for novel proteins from the...

DESCRIPTORS:

...METHODS & EQUIPMENT: DNASTAR, LASERGENE suite component,  
Macintosh compatible, Windows compatible, computer software  
MISCELLANEOUS TERMS: ...Literature Review

7/3,K,AB/3 (Item 2 from file: 55)  
DIALOG(R)File 55:Biosis Previews(R)  
(c) 2003 BIOSIS. All rts. reserv.

0011609864 BIOSIS NO.: 199800404111  
Molecular sequence databases in the field of bioorganic chemistry  
(analytical review)  
AUTHOR: Telezhinskaya I N (Reprint); Ovchinnikova T V  
AUTHOR ADDRESS: M. M. Shemyakin and Yu. A. Ovchinnikov Inst. Bioorg. Chem.,  
Russ. Acad. Sci., ul. Miklukho-Maklaya 16/10, GSP-7, Moscow 117871,  
Russia\*\*Russia  
JOURNAL: Bioorganicheskaya Khimiya 24 (5): p391-400 May, 1998 1998  
MEDIUM: print  
ISSN: 0132-3423  
DOCUMENT TYPE: Article; Literature Review  
RECORD TYPE: Abstract  
LANGUAGE: Russian

ABSTRACT: The main scientific sequence databases of interest for researchers working in the field of bioorganic chemistry are reviewed. Information is given concerning possibilities for rapid access and efficient search for needed information, postal and e-mail addresses, and literature sources in which these databases are comprehensively described.

Molecular sequence databases in the field of bioorganic chemistry  
(analytical review)

DESCRIPTORS:

MISCELLANEOUS TERMS: ...DNASTAR; ...

...Literature Review

7/3,K,AB/4 (Item 3 from file: 55)  
DIALOG(R)File 55:Biosis Previews(R)  
(c) 2003 BIOSIS. All rts. reserv.

**THIS PAGE BLANK (USPTO)**

0008605442 BIOSIS NO.: 199345036423

**DNASTAR** System (LASERGENE for IBM)

BOOK TITLE: **DNASTAR** System (LASERGENE for IBM)

AUTHOR: Dnastar Incorporated (Uk)

BOOK AUTHOR/EDITOR: **DNASTAR**

AUTHOR ADDRESS: 1228 S. Park St., Madison, Wis. 53715, USA\*\*USA  
1992

BOOK PUBLISHER: **DNASTAR** Inc. {a}, 1228 Sotuh Park Street, Madison,  
Wisconsin 53715, USA

DOCUMENT TYPE: Article; Software Review

RECORD TYPE: Abstract

LANGUAGE: English

ABSTRACT: SPECIFICATIONS: IBM or compatible microcomputer. DOS 3.0 or higher. Hard drive with at least 20K free storage capacity. Apple Macintosh microcomputer. Both platforms supported with files interchangeable between systems. Manual included. The software supports most printers and most monitors. Cost: LASERGENE 190, 3000.00; LASERGENE 2000, 4500.00; ENTRY/EDIT-restriction Mapping Package, 750.00.

DESCRIPTION: The **DNASTAR** SYSTEM (LASERGENE) is a comprehensive software package for molecular biologists. The software consists of flexible modular systems with the following functions: DNA analysis, Restriction site analysis, Mapping, Protein analysis, Database searching, Sequence comparison, Shotgun sequencing (gel assembly), Sequence entry and editing, SEQ-EASY II digitizer-talker data entry, and System management. Some of the procedures in the DNA analysis software include sequence display, scanning for patterns, creating a 3 dimensional model of the DNA sequence, and the plotting of codon preference values. The restriction site analysis consists of programs for displaying graphic mini-maps of restriction sites, the scanning of sequence and list restriction sites, and creating a file of restriction enzymes, and other operations. The database search functions contains the program "GENEMAN" which searches "GenBank" or "PIR" for keywords, short sequences or combinations of terms and creates subdatabases and the program "PROSCAN", one purpose of which is to search "PIR" for homologies by the Lipman and Pearson method. Some of the attributes of the sequence comparison programs are to allow the comparison of two DNA sequences, the alignment of DNA sequences and the comparison of two proteins. **DNASTAR** also permits the conversion from other DNA/protein file formats to the **DNASTAR** format. A modem communication program is included. The programs are interactive and menu driven. The menu driven interface is user-friendly and allows quick and easy access to all the software. A demo is available, at no charge, and there is extensive product support.

**DNASTAR** System (LASERGENE for IBM)

BOOK TITLE: **DNASTAR** System (LASERGENE for IBM)

...ABSTRACT: 3000.00; LASERGENE 2000, 4500.00; ENTRY/EDIT-restriction Mapping Package, 750.00. DESCRIPTION: The **DNASTAR** SYSTEM (LASERGENE) is a comprehensive software package for molecular biologists. The software consists of flexible...

...of two DNA sequences, the alignment of DNA sequences and the comparison of two proteins. **DNASTAR** also permits the conversion from other DNA/protein file formats to the **DNASTAR** format. A modem communication program is included. The programs are interactive and menu driven. The...

DESCRIPTORS:

MISCELLANEOUS TERMS: ...Software Review

**THIS PAGE BLANK (USPTO)**

DIALOG(R)File 34:SciSearch(R) Cited Ref Sci  
(c) 2003 Inst for Sci Info. All rts. reserv.

04140670 Genuine Article#: RH341 Number of References: 17  
Title: MACINTOSH SEQUENCE-ANALYSIS SOFTWARE (Abstract Available)  
Author(s): CLEWLEY JP  
Corporate Source: CENT PUBL HLTH LAB,DIV VIRUS REFERENCE/LONDON NW9  
5HT//ENGLAND/  
Journal: MOLECULAR BIOTECHNOLOGY, 1995, V3, N3 (JUN), P221-224  
ISSN: 1073-6085  
Language: ENGLISH Document Type: REVIEW

Abstract: The analysis of information in nucleotide and amino acid sequence data from an investigator's own laboratory, or from the ever-growing worldwide databases, is critically dependent on well planned and written software. Although the most powerful packages previously have been confined to workstations, there has been a dramatic increase over the last few years in the sophistication of the programs available for personal computes, as the speed and power of these have increased. A wide choice of software is available for the Macintosh, including the LaserGene suite of programs from DNASTAR. This review assessed the strengths and weaknesses of LaserGene and concludes that it provides a useful and comprehensive range of sequence analysis tools.

...Abstract: choice of software is available for the Macintosh, including the LaserGene suite of programs from DNASTAR. This review assessed the strengths and weaknesses of LaserGene and concludes that it provides a useful and...

?

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 6.74545 Seconds  
(without alignments)  
2027.556 Million cell updates/sec

Title: US-10-016-768A-1  
Perfect score: 278 KGRTPKXGKRNRYRDSLVE.....RAGSYGVPHSTLEKYKVKER 53  
Sequence: 1 KGRTPKXGKRNRYRDSLVE.....RAGSYGVPHSTLEKYKVKER 53

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: SP\_ARCHAEA:\*
  - 2: SP\_BACTERIA:\*
  - 3: SP\_FUNGI:\*
  - 4: SP\_HUMAN:\*
  - 5: SP\_INVERTEBRATE:\*
  - 6: SP\_MAMMAL:\*
  - 7: SP\_MHC:\*
  - 8: SP\_ORGANELLE:\*
  - 9: SP\_PHAGE:\*
  - 10: SP\_PLANT:\*
  - 11: SP RODENT:\*
  - 12: SP\_VIRUS:\*
  - 13: SP\_VERTEBRATE:\*
  - 14: SP\_UNCLASSIFIED:\*
  - 15: SP\_IVIRUS:\*
  - 16: SP\_BACTERIAP:\*
  - 17: SP\_ARCHAEP:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	278	100.0	1165	5 Q9VUD60	Q9VUD60 drosophila
2	275	98.9	1598	5 Q9SYW8	Q9SYW8 apis mellif
3	217	78.1	185	5 Q22051	Q22051 caenorhabd
4	166	59.7	396	11 Q8C9Q0	Q8C9Q0 mus musculu
5	166	59.7	433	11 Q8BGT2	Q8BGT2 mus musculu
6	166	59.7	572	4 Q96JNO	Q96JNO homo sapien
7	166	59.7	619	4 Q8N3I6	Q8N3I6 homo sapien
8	165	59.4	213	4 Q96NKL	Q96NKL homo sapien
9	165	59.4	517	11 Q8CUG4	Q8CUG4 mus musculu
10	99	35.6	1221	5 Q8MKX3	Q8MKX3 drosophila
11	92.5	33.3	645	5 Q24457	Q24457 drosophila
12	92.5	33.3	660	5 Q9V5N1	Q9V5N1 drosophila
13	92.5	33.3	1064	5 Q9V5N1	Q9V5N1 drosophila
14	92.5	33.3	1085	5 Q24455	Q24455 drosophila
15	84.5	30.4	661	5 Q9V8S2	Q9V8S2 drosophila
16	82	29.5	652	5 Q7168	Q7168 apis mellif

17	70.5	25.4	325	3 Q9VUG7	Q9VUG7 magnaporthe
18	70	25.2	393	11 Q8C9J6	Q8C9J6 mus musculu
19	67.5	24.3	158	17 Q26689	Q26689 methanobac
20	66.5	23.9	663	10 Q04976	Q04976 mangifera
21	64.5	23.2	636	10 Q8LPL0	Q8LPL0 arabidopsis
22	64.5	23.2	728	10 Q9SCV0	Q9SCV0 arabidopsis
23	64.5	23.2	729	10 Q9S2I5	Q9S2I5 arabidopsis
24	64.5	23.0	737	10 Q8L509	Q8L509 cichus sine
25	64	23.0	532	3 Q92205	Q92205 botrytis ci
26	64	23.0	1046	5 Q9W0W2	Q9W0W2 drosophila
27	63.5	22.8	418	16 Q9H544	Q9H544 thizobium a
28	63.5	22.8	721	10 Q9W5J4	Q9W5J4 phaseolus a
29	63.5	22.8	723	10 Q82670	Q82670 cicer ariet
30	63	22.7	368	17 Q9TUG6	Q9TUG6 sulfolobus
31	62.5	22.5	843	10 Q93X58	Q93X58 fragaria an
32	62	22.3	439	10 Q9SDK6	Q9SDK6 oryza sativ
33	61.5	22.1	324	12 Q41274	Q41274 spodoptera
34	61.5	22.1	378	10 Q04529	Q04529 arabidopsis
35	61.5	22.1	378	10 Q93X56	Q93X56 fragaria an
36	61	21.9	722	16 Q8R5T4	Q8R5T4 thermococ
37	61	21.9	739	5 Q8INS2	Q8INS2 drosophila
38	61	21.9	782	5 Q9V1S5	Q9V1S5 saccharopol
39	61	21.8	948	2 Q8KOL9	Q8KOL9 buchnera ap
40	60.5	21.8	730	10 Q9ZPI7	Q9ZPI7 lupinus ang
41	60.5	21.8	731	10 Q9AVS1	Q9AVS1 pyrus pyll
42	60.5	21.8	838	10 Q9ZPI1	Q9ZPI1 lycopersico
43	60.5	21.8	843	10 Q8L3P5	Q8L3P5 oryza sativ
44	60.5	21.8	843	10 Q9AFT0	Q9AFT0 rhizoglia fi
45	60	21.6	100	2 Q9AFT0	Q9AFT0 rhizoglia fi

ALIGNMENTS

RESULT 1

Q9VUD60 ID Q9VUD60 PRELIMINARY: PRT: 1165 AA.

AC Q9VUD60; MEDLINE=20196006; PubMed=10731132;

AD 01-MAY-2000 (TEMBUREL. 13, Created)

DT 01-OCT-2002 (TEMBUREL. 22, Last sequence update)

DT 01-MAR-2003 (TEMBUREL. 23, Last annotation update)

DE CG18389 protein.

DE EIP93F OR CG18389.

GN Drosophila melanogaster (fruit fly)

OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Burkov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieux E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Dou P., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami K.A.,

RA Jaitani W., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

**THIS PAGE IS BLANK**

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulyov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.W., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard D.R., Pui V., Reese M.G.,  
 RA Reimer K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkask R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).

(12)  
 RP SEQUENCE FROM N.A.  
 RA Celinker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Chemp M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorett V., Doup L.E., Doyle C., Drenek D., Farfan D.,  
 RA Ferrera S., Frise E., Galle R.F., Gary N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jaitai M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Paclet J., Fargas V., Park S., Patel S., Pfeiffer B.,  
 RA Phoumenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svrtkask R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 "Sequencing of *Drosophila melanogaster* genome.";  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

(13)  
 RN SEQUENCE FROM N.A.  
 RP Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Jumpy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,  
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Snu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 "Annotation of *Drosophila melanogaster* genome.";  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

(14)  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

(15)  
 RN SEQUENCE FROM N.A.  
 RP Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB003737; AAF55940.3; -  
 DR Flybase: FBgn0013948; E1993F.  
 DR SEQUENCE 1165 AA; 123976 MW; A2556014070BBD8D CRC64;

Query Match 100.0%; Score 278; DB 5; Length 1165;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-24;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 758 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 810

RESULT 2  
 O95YM8 PRELIMINARY; PRT; 1598 AA.  
 AC O95YM8;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created).  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update).  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update).

DE Mdlk-1 protein.  
 GN MBLK-1.  
 OS Apis mellifera (Honeybee).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OC NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21873258; PubMed=11881813;  
 RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,  
 RA Fujiyuki T., Kunieda T., Sekimizu K., Natori S., Kubo T.;  
 RT "Identification of a novel gene, Mdlk-1, that encodes a putative  
 RT transcription factor expressed preferentially in the large-type Kenyon  
 RT cells of the honey bee brain.";  
 RL Insect Mol. Biol. 10:487-494(2001).  
 DR EMBL: AB047034; BAB64330.1; -  
 DR SEQUENCE 1598 AA; 174929 MW; E5475BDD3ACBIEEF CRC64;

Query Match 98.9%; Score 275; DB 5; Length 1598;  
 Best Local Similarity 98.1%; Pred. No. 3.6e-24;  
 Matches 52; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 1031 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 1083

RESULT 3  
 ID Q22051 PRELIMINARY; PRT; 185 AA.  
 AC Q22051;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE TO1C1.3 protein.

GN TO1C1.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lennard N.;  
 RI Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

(12)  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99069613; PubMed=9851916;  
 RX none;  
 RA "Genome sequence of the nematode *C. elegans*: A platform for  
 RT investigating biology.";  
 RT Science 282:2012-2018(1998).  
 DR EMBL: Z68010; CAA92009.1; -  
 DR Wormpep: TO1C1.3; CE03594.  
 DR SEQUENCE 185 AA; 20706 MW; F9F59327B318F641 CRC64;

Query Match 78.1%; Score 217; DB 5; Length 185;  
 Best Local Similarity 73.6%; Pred. No. 2.9e-18;  
 Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

OY 1 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 83 KCTPRKGGKRYNYRDSLSVEAVKAVORGEMSVHRAGSYGVPHSTLEYKVKER 135

RESULT 4  
 O8C9Q0 PRELIMINARY; PRT; 396 AA.  
 AC O8C9Q0;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 12:00:44 ; Search time 6.74545 Seconds  
(without alignment)  
2027.556 Million cell updates/sec

Title: US-10-016-768a-1  
Perfect score: 278  
Sequence: 1 KGTPEKRGKRYNNYDRSLIVE.....RAGSYGVPHSTLEKVKER 53

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Minimum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_ricent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaeal: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278	100.0	1165	5 Q9VD60	Q9VD60 drosophila
2	275	98.9	1598	5 Q95YM8	Q95YM8 apis mellif
3	217	78.1	185	5 Q22051	Q22051 caenorhabdi
4	166	59.7	396	11 Q8C900	Q8C900 mus musculu
5	166	59.7	433	11 Q8BGT2	Q8BGT2 mus musculu
6	166	59.7	572	4 Q96JNO	Q96JNO homo sapien
7	166	59.7	619	4 Q96JL6	Q96JL6 homo sapien
8	165	59.4	213	4 Q96NKL	Q96NKL homo sapien
9	165	59.4	517	11 Q8CJG4	Q8CJG4 mus musculu
10	99	33.6	1221	5 Q24079	Q24079 drosophila
11	92.5	33.3	645	5 Q8MKX3	Q8MKX3 drosophila
12	92.5	33.3	660	5 Q24457	Q24457 drosophila
13	92.5	33.3	1064	5 Q9VSN1	Q9VSN1 drosophila
14	92.5	33.3	1085	5 Q24455	Q24455 drosophila
15	84.5	30.4	661	5 Q9V8S2	Q9V8S2 drosophila
16	82	29.5	652	5 Q77168	Q77168 apis mellif

17	70.5	25.4	325	3 Q9VVG7	Q9VVG7 magnaporthe
18	70	25.2	393	11 Q8C936	Q8C936 mus musculu
19	67.5	24.3	158	17 Q26689	Q26689 methanobact
20	66.5	23.9	663	10 Q04976	Q04976 mangifera
21	64.5	23.2	636	10 Q8LPL0	Q8LPL0 arabidopsis
22	64.5	23.2	728	10 Q9SCV0	Q9SCV0 arabidopsis
23	64.5	23.2	729	10 Q9SZ15	Q9SZ15 arabidopsis
24	64.5	23.2	737	10 Q8LS09	Q8LS09 citrus sine
25	64	23.0	1046	5 Q92205	Q92205 boerhaavia ci
26	64	23.0	1046	5 Q9W0M2	Q9W0M2 drosophila
27	63.5	22.8	418	16 Q98H54	Q98H54 rhizobium
28	63.5	22.8	721	10 Q9M5U4	Q9M5U4 phaseolus a
29	63.5	22.8	723	10 Q82670	Q82670 cicor arret
30	63	22.7	368	17 Q97UG6	Q97UG6 sulfolobus
31	62.5	22.5	843	10 Q93X58	Q93X58 fragaria an
32	62	22.3	439	10 Q9SDK6	Q9SDK6 oryza sativ
33	61.5	22.1	324	12 Q41274	Q41274 spodoptera
34	61.5	22.1	378	10 Q04529	Q04529 arabidopsis
35	61.5	22.1	722	10 Q93X56	Q93X56 fragaria an
36	61	21.9	478	16 Q8R574	Q8R574 thermococ
37	61	21.9	739	5 Q81NS2	Q81NS2 drosophila
38	61	21.9	782	5 Q9V1S5	Q9V1S5 drosophila
39	61	21.9	948	2 Q8KOL9	Q8KOL9 saccharopol
40	60.5	21.8	528	2 Q9KVV9	Q9KVV9 buchera ap
41	60.5	21.8	730	10 Q9ZPI7	Q9ZPI7 lupinus ang
42	60.5	21.8	731	10 Q9AVS1	Q9AVS1 pyrus pyrif
43	60.5	21.8	838	10 Q9ZPI1	Q9ZPI1 lycopersico
44	60.5	21.8	843	10 Q8L3P5	Q8L3P5 oryza sativ
45	60	21.6	100	2 Q9AFT0	Q9AFT0 shigella fl

## ALIGNMENTS

RESULT 1  
ID Q9VD60 PRELIMINARY: PRT: 1165 AA.  
AC Q9VD60: MEDLINE=20196006; PubMed=10731132;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE CG18389 protein.  
DE EIP93F OR CG18389.  
GN Drosophila melanogaster (Fruit fly).  
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkelley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Burdon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,  
RA Burris K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley C., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Foutsier K.J., Evangelista C.C., Ferrar C., Fertiera S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hosten D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K., Nixon K., Nussken D.R., Pacle J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 "The genome sequence of *Drosophila melanogaster*." Science 287:2185-2195(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Bazon J., An H., Baldwin D., Bazon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fartan D.,  
 RA Ferreira S., Friese E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Iobegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Piltman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of *Drosophila melanogaster* genome." Science 287:2185-2195(2000).  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hudecky P., Huang Y., Kaminker J.S., Prochick S.E., Smith C.D.,  
 RA Tudy J.L., Bergman C., Bernan B., Carlson J.W., Ceiniker S.E.,  
 RA Clump M., Drysdale R., Emmert D., Friese E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 "Annotation of *Drosophila melanogaster* genome." Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AE003737; AAF55940.3; -  
 DR FlyBase; FBgn0013948; E393F.  
 SQ SEQUENCE 1165 AA; 123976 MW; A2556014070BEB8D CRC64;  
 Query Match 100.0%; Score 278; DB 5; Length 1165;  
 Best Local Similarity 100.0%; Pred. No. 1.le-24;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 758 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 810  
 RESULT 2  
 ID Q95YV8 PRELIMINARY; PRT; 1598 AA.  
 AC Q95YV8;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE MolK-1 protein.  
 GN MolK-1.  
 OS Apis mellifera (Honeybee).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OX NCBI\_Taxid=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21873258; Pubmed=11881813;  
 RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,  
 RA Fujiyuki T., Kunieda T., Sekimizu K., Natori S., Kubo T.,  
 RT "Identification of a novel gene, MolK-1, that encodes a putative  
 transcription factor expressed preferentially in the large-type Kenyon  
 cells of the honey bee brain." Insect Mol. Biol. 10:487-494(2001).  
 RL EMBL; AB047034; BAB64310.1;  
 DR SEQUENCE 1598 AA; 174929 MW; E5475BD3ACB1EEF CRC64;  
 SQ SEQUENCE 1598 AA; 174929 MW; E5475BD3ACB1EEF CRC64;  
 Query Match 98.9%; Score 275; DB 5; Length 1598;  
 Best Local Similarity 98.1%; Pred. No. 3.6e-24;  
 Matches 52; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 1031 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 1083  
 RESULT 3  
 ID Q22051 PRELIMINARY; PRT; 185 AA.  
 AC Q22051;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE T01C1.3 protein.  
 GN T01C1.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderiinae; Caenorhabditis.  
 OX NCBI\_Taxid=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Leonard N.;  
 RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; Pubmed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C. elegans*: A platform for  
 investigating biology." Science 282:2012-2018(1998).  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z68010; CAA92009.1; -  
 DR WormPep; T01C1.3; CE03594.  
 SQ SEQUENCE 185 AA; 20706 MW; F9F59327B318F641 CRC64;  
 Query Match 78.1%; Score 217; DB 5; Length 185;  
 Best Local Similarity 73.6%; Pred. No. 2.9e-16;  
 Matches 39; Conservative 11; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 83 KGTTPKRGKRYNDRDLSLVEAVKAVRGEMSVHRAGSYGVPHSTLEYKVKER 135  
 RESULT 4  
 ID O8C9Q0 PRELIMINARY; PRT; 396 AA.  
 AC O8C9Q0;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK041621; BAC31007.1; -  
 KW Hypothetical protein.  
 FT NON TER 396 396  
 SQ SEQUENCE 396 AA; 43085 MW; EA4A585F62336E35 CRC64;

Query Match 59.7%; Score 166; DB 11; Length 396;  
 Best Local Similarity 60.4%; Pred. No. 9.9e-12;  
 Matches 32; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 KGTBPKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 337 KQPRKKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 389

## RESULT 5

ID 08BGT2 PRELIMINARY; PRT; 433 AA.  
 AC 08BGT2;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Transcription factor MUR2 (Hypothetical protein).  
 GN MUR2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Kuneda T., Park J., Takeuchi H., Kubo T.;  
 RA "Mus musculus mlr1 and mlr2 mRNA for transcription factor MUR1 and  
 RT MUR2."  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AB076079; BAC20955.1; -  
 DR EMBL; AK041090; BAC30816.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 433 AA; 47124 MW; 73656D1F7E9A041 CRC64;

Query Match 59.7%; Score 166; DB 11; Length 433;  
 Best Local Similarity 60.4%; Pred. No. 1.1e-11;  
 Matches 32; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 KGTBPKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 337 KQPRKKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 389

RESULT 6  
 OY6JUNO PRELIMINARY; PRT; 572 AA.

AC O96JUN0;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Hypothetical protein KIAA1795 (Fragment).  
 GN KIAA1795.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL; AB058698; BAB47424.1; -  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 572 AA; 62730 MW; FB0A401D3F060DF4 CRC64;

Query Match 59.7%; Score 166; DB 4; Length 572;  
 Best Local Similarity 60.4%; Pred. No. 1.5e-11;  
 Matches 32; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 KGTBPKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 476 KQPRKKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 528

## RESULT 7

ID Q8N3J6 PRELIMINARY; PRT; 619 AA.  
 AC Q8N3J6;  
 DT 01-OCT-2002 (TReMBLrel. 22, Created)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFZP451A142.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wamburt R., Heubner D., Mewes H.W., Weil B., Wiemann S.;  
 RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL834245; CAD38921.1; -  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 619 AA; 67378 MW; 791286EC6F8A5110 CRC64;

Query Match 59.7%; Score 166; DB 4; Length 619;  
 Best Local Similarity 60.4%; Pred. No. 1.6e-11;  
 Matches 32; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 KGTBPKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 53  
 Db 523 KQPRKKRGKRYNRYDSDLSVEAVKAVQRGEMSVHRAGSYGVPHSTLEYKVKER 575

## RESULT 8

OY6NKL PRELIMINARY; PRT; 213 AA.  
 ID OY6NKL;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
 DE Hypothetical protein FLJ30696.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Tashiro H., Yamazaki M., Matanabe K., Kumagai A., Itakura S.,  
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Ohtsuki T., Sato H., Ota T., Makamatsu A., Ichih S., Yamamoto J.,  
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
 RA Magatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,  
 RA Masuho Y., Nagai K., Isogai T.,  
 RT "NBD human cDNA sequencing project."  
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AK055258; BAB70892.1;  
 SO SEQUENCE 213 AA; 23477 MW; 4D7F6CABF95251B2 CRC64;  
 Query Match 59.4%; Score 165; DB 4; Length 213;  
 Best Local Similarity 60.4%; Pred. No. 6.4e-12;  
 Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 KGTTPKRGKRYNYPDRSLVEAVKAVQSGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 124 KQPKKRGKRYNYPDRSLVEAVKAVQSGEMSVHRAGSYGVPHSTLEYKVKER 176  
 RESULT 9  
 OX OX8CJG4 PRELIMINARY; PRT; 517 AA.  
 AC OX8CJG4;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Transcription factor MLR1.  
 GN MLR1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Kunieda T., Park J., Takeuchi H., Kubo T.,  
 RT "Mus musculus mlr1 and mlr2 mRNA for transcription factor MLR1 and  
 RT MLR2."  
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AB076078; BAC20954.1;  
 SO SEQUENCE 517 AA; 57316 MW; C97403D3D296C52E CRC64;  
 Query Match 59.4%; Score 165; DB 11; Length 517;  
 Best Local Similarity 60.4%; Pred. No. 1.8e-11;  
 Matches 32; Conservative 6; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 KGTTPKRGKRYNYPDRSLVEAVKAVQSGEMSVHRAGSYGVPHSTLEYKVKER 53  
 DB 429 KQPKKRGKRYNYPDRSLVEAVKAVQSGEMSVHRAGSYGVPHSTLEYKVKER 481  
 RESULT 10  
 ID OX24079 PRELIMINARY; PRT; 1221 AA.  
 AC OX24079;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Ecdysone-regulated (E93).  
 GN EIP93F OR E93 OR CG18389.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON S;  
 RX MEDLINE=96018744; PubMed=7556910;  
 RA Baehrcke E.H., Thummel C.S.;  
 RT "The Drosophila E93 gene from the 93F early puff displays stage- and  
 RT tissue-specific regulation by 20-hydroxyecdysone.";  
 RL Dev. Biol. 171:85-97(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON S;  
 RX MEDLINE=95042758; PubMed=7954827;  
 RA Woodward C.T., Baehrcke E.H., Thummel C.S.;  
 RT "A molecular mechanism for the stage specificity of the Drosophila  
 RT prepupal genetic response to ecdysone.";  
 RL Cell 79:607-615(1994).  
 DR EMBL; U25686; AAA83228.1;  
 DR FLYbase; FBgn0013948; EIP93F.  
 SO SEQUENCE 1221 AA; 13175 MW; F949BF637EB377B8 CRC64;  
 Query Match 35.6%; Score 99; DB 5; Length 1221;  
 Best Local Similarity 100.0%; Pred. No. 0.0043;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KGTTPKRGKRYNYPDRSL 18  
 DB 758 KGTTPKRGKRYNYPDRSL 775  
 RESULT 11  
 OX OX8MKX3 PRELIMINARY; PRT; 645 AA.  
 AC OX8MKX3;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE CG2368-PD.  
 GN PSD OR CG2368.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brooksstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Canter A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ijzerman C.,  
 RA Jalaali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Sieden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spiedling A.C., Stapleton M., Strong M., Sun E.,  
 Svitskas R., Tector C., Turner R., Venter G., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RA Celisner S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,  
 Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 Dodson K., Dorsett V., Doup L.E., Doyle C., Dresek D., Farfan D.,  
 Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
 Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 Pacleb J., Paragias V., Park S., Patel S., Pfeiffer B.,  
 Phanaphavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svitskas R., Tector C., Tyler D.,  
 Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celisner S.E.,  
 Krommiller B., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seale S.M.J., Smith E., Shu S., Smurniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RA EMBL: AE003829; AAM68770.1; -;  
 DR FlyBase; FBgn004399; psg.  
 DR InterPro; IPR002197; HTH\_Fls.  
 DR TIGRFAMs; TIGR01199; HTH\_Fls; 1.  
 SO SEQUENCE 645 AA; 70298 MW; 4872F47175060529 CRC64;  
 Query Match 33.3%; Score 92.5; DB 5; Length 645;  
 Best Local Similarity 35.3%; Pred. No. 0.013; Indels 1; Gaps 1;  
 Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
 DB 352 TPKEGKGRKSNMEDALQNALTEALRSGQISANKASAKAFIPSTL-YKTIARR 401  
 RESULT 12  
 ID 024457 PRELIMINARY; PRT; 660 AA.  
 AC 024457;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE PIPEQUEAK protein (ORF-B).

GN PSQ OR CG2368.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96134923; PubMed=8557044;  
 RA Weber U., Siegel V., Mlodzik M.;  
 RT "Pipequeak encodes a novel nuclear protein required downstream of  
 RT seven-up for the development of photoreceptors R3 and R4.";  
 RL EMO J. 14:6247-6257 (1995).  
 DR EMBL: X90986; CAA62475.1; -;  
 DR FlyBase; FBgn004399; psg.  
 DR InterPro; IPR002197; HTH\_Fls.  
 DR TIGRFAMs; TIGR01199; HTH\_Fls; 2.  
 SO SEQUENCE 660 AA; 71818 MW; 6E251F440326547F CRC64;  
 Query Match 33.3%; Score 92.5; DB 5; Length 660;  
 Best Local Similarity 35.3%; Pred. No. 0.013; Indels 1; Gaps 1;  
 Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
 DB 346 TPKEGKGRKSNMEDALQNALTEALRSGQISANKASAKAFIPSTL-YKTIARR 395  
 RESULT 13  
 ID 09V5N1 PRELIMINARY; PRT; 1064 AA.  
 AC 09V5N1; 09V5N2; 024184; 024187;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE psg protein (IDJ3470p).  
 GN PSQ OR CG2368.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM A).  
 RC TISSUE=Ovary;  
 RX MEDLINE=95220671; PubMed=7705633;  
 RA Horowitz H., Berg C.A.;  
 RT "Aberrant splicing and transcription termination caused by p element  
 RT insertion into the intron of a *Drosophila* gene.";  
 RL Genetics 139:327-335 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B).  
 RC TISSUE=Ovary;  
 RX MEDLINE=96223200; PubMed=8674425;  
 RA Horowitz H., Berg C.A.;  
 RT "The *Drosophila* pipequeak gene encodes a nuclear BTB-domain-containing  
 RT protein required early in oogenesis.";  
 RL Development 122:1859-1871 (1996).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS A AND 2).  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
 RA Abtil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glaser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
 RA Jajali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Paclob J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [4]  
 RP "SEQUENCE FROM N.A."  
 RC STRAIN=Berkeley;  
 RA Champel M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuncio J., Paclob J., Paragas V., Park S., Phouenavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,  
 RL Submitted (DGC-2001) to the EMBL/Genbank/DBJ databases.  
 CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A/1 (SHOWN HERE), B AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 DR EMBL; U48358; AAC47153.1; -;  
 DR EMBL; U48402; AAC47154.1; -;  
 DR EMBL; AE003829; AAF58769.1; -;  
 DR EMBL; AE003829; AAF58770.1; -;  
 DR EMBL; AY069588; AAL39733.1; -;  
 DR FLYBase; FBgn0004399; psq.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR002197; HTH\_Fis.  
 DR Pfam; PF00651; BTB; 1.  
 DR SMART; SM00225; BTB; 1.  
 DR TIGRfams; TIGR01199; HTH\_fis; 2.  
 DR PROSITE; PS50097; BTB; 1.  
 KM Alternative splicing.  
 FT VARSPLIC 1 429 MISSING (IN ISOFORM B).  
 FT VARSPLIC 719 736 MISSING (IN ISOFORM 2).  
 FT CONFLICT 1020 1020 Q -> QQ (IN REF. 1 AND 2).  
 SQ SEQUENCE 1064 AA; 114984 MW; 77420C782DE6ECAS CRC64;  
 Query Match 33.3%; Score 92.5; DB 5; Length 1064;  
 Best Local Similarity 35.3%; Pred. No. 0.022;  
 Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
 Oy 3 TRPKRGKRYNVDRLSLVEAVKAVGEMSVHAGSYGVPHSTLEYKVER 53  
 Db 771 TPKEGGKTSWNEALQNALALRLRGQISANKAKAFIPSTL-YKIAAR 820  
 RESULT 14.  
 ID 024455 PRELIMINARY; PRT; 1085 AA.  
 AC 024455; 024456; 024003;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE Pipsqueak protein (BTB-V protein domain).

GN PSQ OR CG2368.  
 OS *Drosophila melanogaster* (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
 RX MEDLINE=96134923; PubMed=8557044;  
 RA Weber U., Siegel V., Mlodzik M.,  
 RT "Pipsqueak encodes a novel nuclear protein required downstream of  
 RT seven-up for the development of photoreceptors R3 and R4.";  
 RL EMO J. 14:6247-6257(1995).  
 RN [2]  
 RP SEQUENCE OF 8-105 FROM N.A.  
 RX MEDLINE=95024186; PubMed=7938017;  
 RA Zollman S., Godt D., Prive G.G., Coudere J.L., Lasaki F.A.;  
 RT "The BTB domain, found primarily in zinc finger proteins, defines an  
 RT evolutionarily conserved family that includes several developmentally  
 RT regulated genes in *Drosophila*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).  
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 DR EMBL; X90986; CAA62473.1; -;  
 DR EMBL; X90986; CAA62474.1; -;  
 DR EMBL; U14402; AAA50837.1; -;  
 DR FLYBase; FBgn0004399; psq.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR002197; HTH\_Fis.  
 DR Pfam; PF00651; BTB; 1.  
 DR SMART; SM00225; BTB; 1.  
 DR TIGRfams; TIGR01199; HTH\_fis; 2.  
 DR PROSITE; PS50097; BTB; 1.  
 KM Alternative splicing.  
 FT VARSPLIC 428 535  
 FT VARSPLIC 536 1084  
 FT SEQUENCE 1085 AA; 117039 MW; EF32BFC73C2B737D CRC64;  
 Query Match 33.3%; Score 92.5; DB 5; Length 1085;  
 Best Local Similarity 35.3%; Pred. No. 0.023;  
 Matches 18; Conservative 16; Mismatches 16; Indels 1; Gaps 1;  
 Oy 3 TRPKRGKRYNVDRLSLVEAVKAVGEMSVHAGSYGVPHSTLEYKVER 53  
 Db 771 TPKEGGKTSWNEALQNALALRLRGQISANKAKAFIPSTL-YKIAAR 820  
 RESULT 15  
 ID 09V852 PRELIMINARY; PRT; 661 AA.  
 AC 09V852;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE CG7230 protein (RIBBON).  
 GN RIB OR CG7230.  
 OS *Drosophila melanogaster* (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731133;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhattacharya D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milbina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskeas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Shim K., Blake K.J., Jack J., Kraanow M.A.;  
 RT "The Drosophila ribbon gene encodes a nuclear BTB domain protein that  
 RT promotes epithelial migration and morphogenesis.";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003796; AAF57588.1; -;  
 DR EMBL; AF416603; AAL1905.1; -;  
 DR FLYBase; FBgn003254; rib.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR InterPro; IPR002197; HTR\_Fls.  
 DR Pfam; PF00651; BTB; 1.  
 DR SMART; SM00225; BTB; 1.  
 DR TIGRFAMs; TIGR01199; HTR\_Fls; 1.  
 DR PROSITE; PS50097; BTB; 1;  
 DR SEQUENCE 661 AA; 70977 MW; 9A827146FCF1122E CRC64;

Query Match 30.4%; Score 84.5; DB 5; Length 661;  
 Best Local Similarity 37.3%; Pred. No. 0.12;  
 Matches 19; Conservative 13; Mismatches 18; Indels 1; Gaps 1;  
 QY 2 GTRPRKGYRNRDRLSLVAVKAVQRCMSYHRAAGSYGVPHSTLEKYKE 52  
 Db 361 GKPEWKRYKQYTRADWCAIQAVREG-MSALQASRKYGLPRTLYDKVRK 410

Search completed: October 28, 2003, 12:04:21  
 Job time : 12.7455 secs

**THIS PAGE BLANK**

**From:** Davis, Minh-Tam  
**Sent:** Tuesday, October 28, 2003 1:26 PM  
**To:** STIC-ILL  
**Subject:** Reprint request for 10/016768

1) Apoptosis induced by topoisomerase inhibitors.

Sordet Olivier; Khan Qasim A; Kohn Kurt W; Pommier Yves  
Laboratory of Molecular Pharmacology, Center for Cancer Research,  
National Cancer Institute, NIH, Bethesda, Maryland 20892-4255, USA.  
Curr Med Chem Anti-Canc Agents (Netherlands) Jul 2003, 3 (4) p271-90  
ISSN 1568-0118 Journal Code: 101123597  
Document type: Journal Article; Review; Review, Academic  
Languages: ENGLISH

2) 15271751- 22761780 PMID: 12879973

Cerebellar granule cells as a model to study mechanisms of neuronal  
apoptosis or survival in vivo and in vitro.

Contestabile Antonio  
Department of Biology, University of Bologna, Italy.  
acontest@alma.unibo.it  
Cerebellum (England) Jan-Mar 2002, 1 (1) p41-55, ISSN 1473-4222  
Journal Code: 101089443  
Document type: Journal Article; Review; Review, Academic

3) IALOG(R)File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

15498572 22879252 PMID: 14517806

Apoptosis versus oncotic necrosis in hepatic ischemia/reperfusion  
injury.

Jaeschke Hartmut; Lemasters John J  
Liver Research Institute, University of Arizona, College of Medicine,  
Room 6309, 1501 N. Campbell Avenue, Tucson, Arizona, USA.  
Jaeschke@email.arizona.edu  
Gastroenterology (United States) Oct 2003, 125 (4) p1246-57; ISSN  
0016-5085 Journal Code: 0374630  
Contract/Grant No.: AA12916; AA; NIAAA; AG13637; AG; NIA; DK37034; DK;  
NIDDK; DK59340; DK; NIDDK; ES06091; ES; NIEHS

4) Diversity in the mechanisms of neuronal cell death.

Yuan Junying; Lipinski Marta; Degterev Alexei  
Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue,  
02115, Boston, MA, USA  
Neuron (United States) Oct 9 2003, 40 (2) p401-13, ISSN 0896-6273  
Journal Code: 8809320

Thank you.

MINH TAM DAVIS  
ART UNIT 1642, ROOM 8A01, MB 8E12  
305-2008

**THIS PAGE BLANK (USPTO)**

# Apoptosis Versus Oncotic Necrosis in Hepatic Ischemia/Reperfusion Injury

HARTMUT JAECHKE\* and JOHN J. LEMASTERS\*

\*Liver Research Institute, University of Arizona, Tucson, Arizona; and \*Department of Cell and Developmental Biology, University of North Carolina, Chapel Hill, North Carolina

Warm and cold hepatic ischemia followed by reperfusion leads to necrotic cell death (oncosis), which often occurs within minutes of reperfusion. Recent studies also suggest a large component of apoptosis after ischemia/reperfusion. Here, we review the mechanisms underlying adenosine triphosphate depletion-dependent oncotic necrosis and caspase-dependent apoptosis, with emphasis on shared features and pathways. Although apoptosis causes internucleosomal DNA degradation that can be detected by terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick-end labeling and related assays, DNA degradation also occurs after oncotic necrosis and leads to pervasive terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick-end labeling staining far in excess of that for apoptosis. Similarly, although apoptosis can occur in a physiological setting without inflammation, in pathophysiological settings apoptosis frequently induces inflammation because of the onset of secondary necrosis and stimulation of cytokine and chemokine formation. In liver, the mitochondrial permeability transition represents a shared pathway that leads to both oncotic necrosis and apoptosis. When the mitochondrial permeability transition causes severe adenosine triphosphate depletion, plasma membrane failure and necrosis ensue. If adenosine triphosphate is preserved, at least in part, cytochrome c release after the mitochondrial permeability transition activates caspase-dependent apoptosis. Mitochondrial permeability transition-dependent cell death illustrates the concept of *necrapoptosis*, whereby common pathways lead to both necrosis and apoptosis. In conclusion, oncotic necrosis and apoptosis can share features and mechanisms, which sometimes makes discrimination between the 2 forms of cell death difficult. However, elucidation of critical cell death pathways under clinically relevant conditions will show potentially important therapeutic intervention strategies in hepatic ischemia/reperfusion injury.

**H**epatic ischemia/reperfusion injury occurs in diverse circumstances, including liver surgery (e.g., a Pringle maneuver during tumor resection or liver trauma),

liver preservation for transplantation, veno-occlusive disease, hemorrhagic shock-resuscitation, and heart failure. Different injury mechanisms contribute to the overall pathophysiology of hepatic ischemia/reperfusion injury.<sup>1-6</sup> Although ischemic stress itself primes cells for damage and will eventually cause cell death, cell injury often does not manifest itself until after the ischemic liver is reperfused.<sup>7</sup>

Production of reactive oxygen species, including superoxide, hydrogen peroxide, and hydroxyl radicals, has long been implicated in reperfusion injury, but oxygen-independent factors are important as well, such as tissue pH changes during ischemia/reperfusion.<sup>8</sup> Inflammatory responses<sup>2,6</sup> and microcirculatory problems<sup>4</sup> further aggravate injury after reperfusion. Ischemia/reperfusion activates Kupffer cells, the resident macrophages of the liver, and functional inactivation of Kupffer cells attenuates injury during early and late reperfusion.<sup>9-13</sup> Kupffer cells activated after reperfusion generate reactive oxygen species, proinflammatory cytokines, chemokines, and other mediators that contribute to postischemic tissue injury and to the systemic inflammatory response syndrome and multiorgan failure that may follow a severe ischemic insult to the liver.<sup>14</sup> Together with activated complement factors,<sup>15</sup> these inflammatory mediators activate and recruit neutrophils into the postischemic liver,<sup>16,17</sup> which generates even more reactive oxygen<sup>18,19</sup> and releases additional proteases and other degradative enzymes.<sup>20</sup> In addition to the inflammatory response, vasoconstriction of sinusoids induced by endothelin-1<sup>21</sup> promotes heterogeneous closure of many microvessels, which prolongs ischemia in certain areas of the liver even after reperfusion.<sup>22</sup>

*Abbreviations used in this paper:* IκB, inhibitor of nuclear factor κB; MPT, mitochondrial permeability transition; NFκB, nuclear factor-κB; TNF, tumor necrosis factor; TUNEL, terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick-end labeling.

© 2003 by the American Gastroenterological Association

0016-5085/03/\$30.00

doi:10.1053/S0016-5085(03)01209-5

**THIS PAGE BLANK (USPTO)**

Most of the described mechanisms of reperfusion injury generally assume cell damage that involves oncotic necrosis. However, several recent reports propose that apoptosis occurs in postischemic heart, liver, and other tissues.<sup>23-25</sup> Postischemic apoptosis would seem to contradict earlier findings of necrotic cell death. Thus, confusion and uncertainty exist concerning the actual mode of cell killing after ischemia/reperfusion. Accordingly, the goal of this overview is to discuss the distinctions between apoptosis and necrosis and to evaluate critically the methods and approaches used to quantify apoptotic and necrotic cell death to reach conclusions regarding the pathophysiological role of each mode of cell death in hepatic ischemia/reperfusion injury.

### **Oncotic Necrosis (Oncosis) in Ischemia/Reperfusion Injury**

The primary stress in ischemia to liver and most other solid tissues is loss of mitochondrial adenosine triphosphate (ATP) production. The resulting ATP depletion leads to cellular swelling, rounding and swelling of mitochondria, dilatation of the endoplasmic reticulum, and formation of plasma membrane protrusions called *blebs*.<sup>26,27</sup> Blebs are a consequence of ATP depletion and likely represent a response to disrupted cellular volume control and cytoskeletal disturbances. After briefer periods of ischemia/anoxia, bleb formation rapidly reverses after reoxygenation, but necrotic cell death occurs after longer periods. Just before cell death, hepatocytes and hepatic sinusoidal cells develop a metastable state, which is characterized by mitochondrial permeabilization, lysosomal disruption, bleb coalescence and growth, cell swelling, and leakage of anionic solutes.<sup>28-30</sup> Opening of glycine-sensitive anion channels that conduct chloride and various organic anions initiates the metastable state and a phase of rapid colloid osmotic swelling.<sup>30</sup> Cell death then occurs by failure of the plasma membrane permeability barrier, often caused by bleb rupture.<sup>28,29</sup> Plasma membrane permeabilization causes release of cellular enzymes and other contents, labeling with vital dyes such as trypan blue, and development of histological changes known as *necrosis*. The release of cellular contents also initiates an inflammatory response during reperfusion. Over time, macrophages gradually resorb the remnants of the necrotic tissue, which is replaced by scar tissue. Taken together, the observations of postischemic cell swelling, vacuolation, karyolysis, and cell content release, affecting cells in large areas of the tissue in combination with a substantial inflammatory response, are characteristic features of a necrotic cell

death process, more recently renamed *oncosis* or *oncotic necrosis*.<sup>31</sup>

### **Morphological Features of Apoptosis**

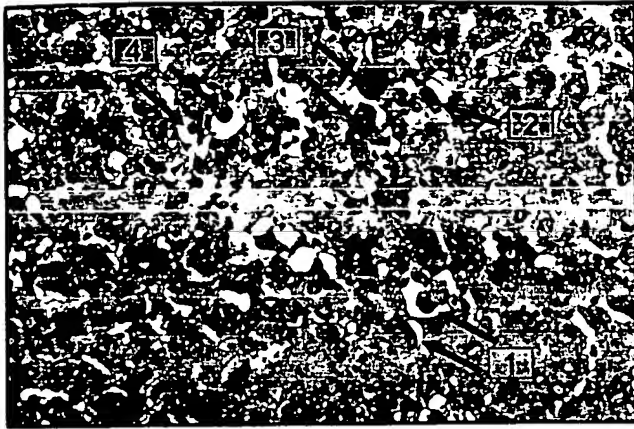
The original description of apoptotic cell death was based on morphology.<sup>32</sup> The classic morphological features of apoptosis include cellular shrinkage, nuclear condensation, chromatin margination, and fragmentation of both the nucleus and cytoplasm into apoptotic bodies, which are phagocytosed and degraded by phagocytes, neighboring cells, or both (Figure 1). The original definition of apoptosis describes the cytoplasmic organelles of apoptotic cells as remaining normal in appearance, in marked contrast to necrosis, although many more recent studies show mitochondrial swelling, changes to the endoplasmic reticulum, increased autophagy, and other cytoplasmic changes during apoptosis.<sup>33</sup> In classic apoptotic cell death, intracellular contents are not released, and a consequent inflammatory response fails to develop. Functionally, apoptosis eliminates excess and unneeded cells during development and damaged and worn-out cells during normal tissue turnover. Characteristically, apoptosis affects individual isolated cells in a tissue, rather than groups of contiguous cells. Under certain conditions, the apoptotic cell death program may not go to completion. Instead, secondary necrosis supervenes, resulting in the release of proinflammatory intracellular contents.<sup>34</sup>

### **Signaling Mechanisms in Hepatocellular Apoptosis**

We first briefly review some basic background information on apoptotic signaling pathways in hepatocytes to place into context the discussion of whether postischemic cell death is caused by apoptosis. During the last decade, dramatic progress has been made in the elucidation of the intracellular signaling mechanisms of apoptosis.<sup>35-39</sup> A variety of mediators, including tumor necrosis factor (TNF)- $\alpha$ , Fas ligand, and tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), activate a so-called extrinsic pathway to apoptosis. As illustrated for TNF- $\alpha$  in Figure 2, these proapoptotic mediators first bind to their respective receptors, which cause receptor oligomerization and the association of various adapter proteins, including Fas-associated death domain, TNF- $\alpha$  receptor-associated death domain, and TNF- $\alpha$  receptor-associated factor. Fas-associated death domain and TNF- $\alpha$  receptor-associated death domain promote binding of procaspase 8 and its proteolytic activation to catalytic caspase 8. If sufficient amounts of caspase 8 are

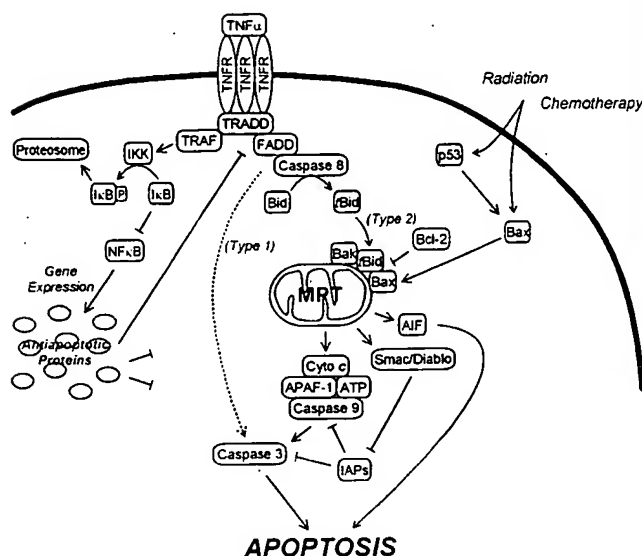
**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



**Figure 1.** Liver histology of galactosamine-induced apoptosis. Characteristic morphology is shown of rat hepatocytes undergoing apoptotic cell death 6 hours after treatment with galactosamine (500 mg/kg). Features of apoptosis include cell shrinkage (1), chromatin margination (2), chromatin condensation and fragmentation (3), and formation of apoptotic bodies (4).

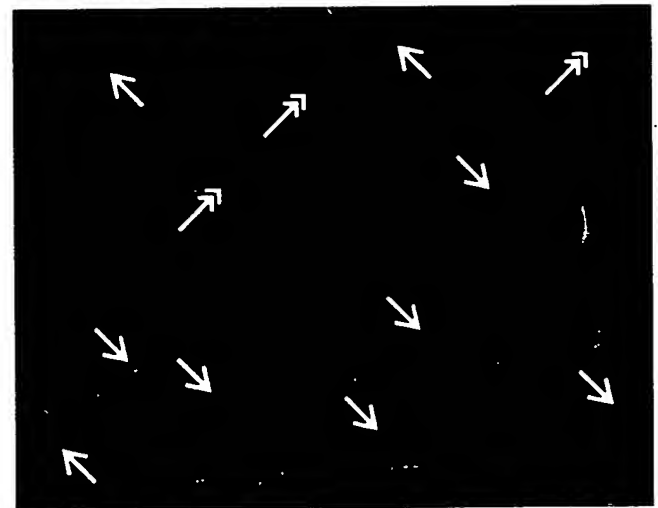
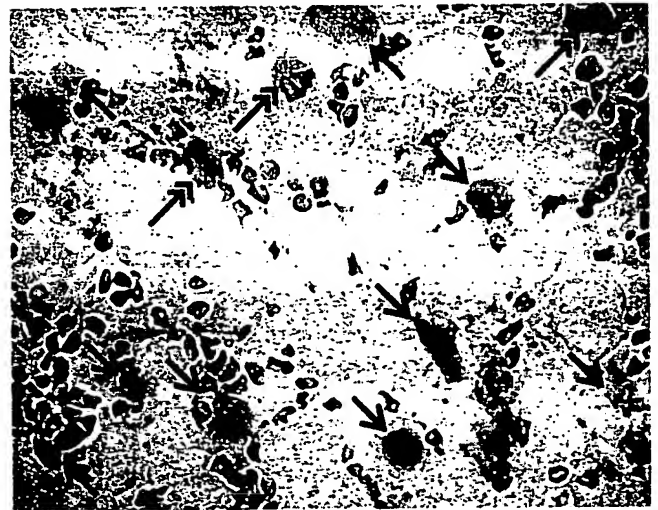
generated at the receptor, caspase 8 can directly activate procaspase 3 (type 1 pathway).<sup>40</sup> However, in hepatocytes, the receptor signal needs to be amplified through mitochondria (type 2 pathway).<sup>40,41</sup> Caspase 8 cleaves Bid, a BH3 domain-only Bcl-2 family member, to an active fragment,  $\ell$ Bid, which translocates to mitochondria.  $\ell$ Bid translocation leads to release of soluble proteins from mitochondria that activate caspases and initiate apoptotic nuclear changes.<sup>41</sup> These protein factors include cytochrome *c*, apoptosis-inducing factor, and Smac/Diablo, which reside in the intermembrane space



**Figure 2.** Scheme of apoptotic signaling in hepatocytes. AIF, apoptosis-inducing factor; APAF-1, apoptosis-activating factor-1; IKK, I $\kappa$ B kinase; TRADD, tumor necrosis factor- $\alpha$  receptor-associated death domain; TRAF, tumor necrosis factor- $\alpha$  receptor-associated factor.

between the mitochondrial inner and outer membranes.<sup>42-45</sup>

The mechanisms that induce the release of mitochondrial intermembrane proteins remain controversial. In hepatocytes, TNF- $\alpha$ - and Fas-dependent signaling induces the onset of the mitochondrial permeability transition (MPT). The MPT occurs from the opening of a pore, the permeability transition pore, which is highly conductive to solutes of molecular weight up to approximately 1500 daltons.<sup>46</sup> As a consequence of permeability transition pore opening, mitochondria depolarize, uncouple, and undergo large amplitude swelling. This



**Figure 3.** Trypan blue staining and TUNEL labeling after cold liver storage and orthotopic rat liver transplantation. A rat liver was stored in cold University of Wisconsin solution for 44 hours and transplanted into a recipient rat. After 15 minutes of implantation, the liver graft was infused with trypan blue and fixed. The upper panel shows trypan blue uptake into the nuclei of nonviable cells, indicating oncotic necrosis, whereas the lower panel shows TUNEL-positive nuclei labeled with green fluorescence, indicating DNA strand breaks. Note that all TUNEL-positive cells stain with trypan blue (arrows), whereas some trypan blue-labeled cells stain weakly or not at all with TUNEL (double arrows). (X.-X. Peng, et al., unpublished data, January 2003).

**THIS PAGE BLANK (USPTO)**

swelling causes a rupture of the outer membrane and a release of proteins from the intermembrane space. Other mechanisms for cytochrome *c* release also seem to exist. In some models, *t*Bid interaction with either Bax or Bak, 2 other proapoptotic members of the Bcl-2 family, forms channels in the mitochondrial outer membrane that release cytochrome *c* and a number of other, larger proteins from the intermembrane space. The nature and composition of these channels, however, remain poorly understood.<sup>43,47,48</sup> Bcl-2 and other antiapoptotic Bcl-2 family members block cytochrome *c* release. The mechanism for the antiapoptotic action of Bcl-2 may involve blockade of the MPT and/or antagonism of Bax/Bak-dependent pore formation in the mitochondrial outer membrane.

After its release from mitochondria, cytochrome *c* forms a complex with apoptosis-activating factor-1, ATP (or deoxyadenosine triphosphate), and procaspase 9 (Figure 2). This complex, called an *apoptosome*, proteolytically activates caspase 9, which in turn activates procaspase 3 to caspase 3.<sup>49</sup> Through action on a variety of substrates, caspase 3 activity initiates the final execution stages of apoptosis, including cell shrinkage, surface blebbing, internucleosomal DNA hydrolysis, phosphatidyl serine externalization on the plasma membrane, chromatin margination, and nuclear lobulation.

In general, the type 2 apoptotic signaling pathway through the mitochondria, as it occurs in hepatocytes, is faster than the type 1 pathway and can be better regulated. However, if the type 2 pathway is blocked by inhibition of the MPT with cyclosporin A, caspase 3 activation and apoptosis will still occur, but at a slower rate, via a type 1 pathway, but without mitochondrial depolarization, the MPT, or cytochrome *c* release (Figure 2).<sup>50</sup> Adding to this redundancy is the so-called intrinsic

pathway to apoptosis<sup>51</sup> (Figure 2). Such pathways may or may not involve p53-dependent gene expression but may activate apoptosis by still incompletely understood mechanisms through translocation of Bax and other proapoptotic Bcl-2 family members to the mitochondria to cause cytochrome *c* release and the activation of caspases 9 and 3.

Death receptors also initiate survival signals. For example, ligation and oligomerization of TNF- $\alpha$  receptors and Fas promote receptor association of another adapter protein, TNF- $\alpha$  receptor-associated factor. TNF- $\alpha$  receptor-associated factor in turn initiates inhibitor of nuclear factor- $\kappa$ B (I $\kappa$ B) kinase activation, I $\kappa$ B phosphorylation, and subsequent degradation of I $\kappa$ B in proteasomes. The disappearance of I $\kappa$ B de-represses nuclear factor- $\kappa$ B (NF $\kappa$ B), which translocates to the nucleus to induce expression of several antiapoptotic genes that prevent apoptosis from occurring, including inhibitor-of-apoptosis proteins (Figure 2).<sup>52,53</sup> Survival signaling through NF $\kappa$ B is so strong that to induce apoptosis in cultured hepatocytes, NF $\kappa$ B-dependent gene expression must be blocked by using protein or messenger RNA synthesis inhibitors or by expressing an I $\kappa$ B superrepressor that has been mutated to lack a phosphorylation site for I $\kappa$ B kinase.<sup>54</sup>

## Assessment of Apoptotic Cell Death

As a result of an increasing understanding of the mechanisms and pathways to apoptotic cell death, more and more biochemical and immunologic assays are being developed and used to characterize apoptosis (Table 1). Today, apoptosis can be monitored in vitro and in vivo

**Table 1.** Assays for Apoptosis

Nuclear morphology (chromatin condensation and nuclear lobulation/fragmentation) in histological sections and after fluorescent staining with DAPI, propidium iodide, and so on
Internucleosomal DNA cleavage (DNA ladder after starch gel electrophoresis; ELISA for DNA fragments)
TUNEL and related assays (in situ detection of double-stranded DNA breaks)
Annexin V (phosphatidyl serine externalization)
Caspase assays (especially caspases 3, 2, 8, and 9)
Enzyme assays using fluorogenic substrates
Immunocytochemistry with specific antibodies against activated caspases
Western blotting to show a decrease of the proenzyme and the appearance of active fragments
Cleavage of caspase substrates (e.g., PARP cleavage)
Nuclear staining with supravital dyes for secondary necrosis (trypan blue; propidium iodide)
Mitochondrial depolarization assessed with potential-indicating fluorophores (rhodamine 123, tetramethylrhodamine methylester, JC-1, and so on)
Cytochrome <i>c</i> release into cytosol (Western blot; immunocytochemistry)
Translocation of proapoptotic proteins (Bax, Bid, and so on) to mitochondria; proteolytic cleavage of Bid (Western blot; immunocytochemistry)

DAPI, 4',6-diamidino-2-phenylindole; ELISA, enzyme-linked immunosorbent assay; JC-1, 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolyl carboxycyanine iodide. PARP, poly (adenosine diphosphate-ribose) polymerase.

**THIS PAGE BLANK (USPTO)**

by such diverse techniques as enzyme assays for activated caspases, Western blot analyses for caspase processing; annexin V labeling for phosphatidyl serine externalization, cleavage of poly (adenosine diphosphate-ribose) polymerase and other targets of caspase 3 proteolytic action, trypan blue and propidium iodide staining, and mitochondrial depolarization and cytochrome *c* release; in addition to classic morphological criteria. In particular, a distinctive form of random DNA cleavage between nucleosomes occurs in apoptosis, which produces DNA fragments in multiples of approximately 190 base pairs (the length of DNA from 1 nucleosome to the next). This pattern of DNA cleavage produces a characteristic ladder pattern after gel electrophoresis. Other methods to assess DNA cleavage are by enzyme-linked immunosorbent assay kits that detect DNA fragments and by the terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick-end labeling (TUNEL) assay (Table 1).

It is important to note that apoptosis represents a constellation of events, and no single change is necessarily a required event in apoptosis or is unique to apoptosis. For example, mitochondrial depolarization, swelling, and cytochrome *c* release also typically occur in oncotic necrosis, and the trypan blue and propidium iodide labeling of so-called late apoptosis actually represents a phenomenon of secondary necrosis associated with loss of plasma membrane integrity. Necrotic cell death also leads to annexin V labeling,<sup>55</sup> because after lysis of the plasma membrane, annexin V gains entrance to the interior of cells and the internal surface of the plasma membrane, where phosphatidylserine normally resides. Caspase 3 activation is perhaps most uniquely associated with apoptosis, but not all forms of apoptosis require caspase 3 activation.

Necrosis also causes DNA cleavage, although such cleavage is not characteristically internucleosomal. On gel electrophoresis, such cleavage leads to a smear of many different molecular weight fragments rather than a ladder pattern of multiples of 190 base pairs. However, DNA fragmentation with a ladder pattern was also reported during necrosis.<sup>56,57</sup> Calcium-dependent activation of endonucleases may be responsible for this effect.<sup>58</sup> Techniques such as the TUNEL assay may not distinguish the internucleosomal DNA cleavage of apoptosis from the much less ordered DNA cleavage of necrosis. In liver and other tissues, TUNEL labeling occurs during ischemic necrosis and after toxicant-induced necrotic cell killing.<sup>59-61</sup> Indeed, after reperfusion of livers stored for transplantation, the same cells showing TUNEL labeling, a presumptive indicator of apoptosis, also labeled

with trypan blue, an indicator of oncotic necrosis, whereas not all trypan blue-stained cells labeled with TUNEL (Figure 3). These observations are consistent with postnecrotic DNA hydrolysis as the basis for TUNEL conversion in necrotic tissue, as previously proposed.<sup>59-61</sup>

Overall, the most reliable method to identify apoptotic cell death is morphology. Vital dyes and simple histology can readily visualize nuclear morphology (chromatin condensation and fragmentation). Once apoptosis as the mode of cell death is established by the characteristic morphological changes, any number of other parameters listed in Table 1 can be used to further support the hypothesis and delineate specific signaling pathways.

### Apoptotic Cell Death During Hepatic Ischemia/Reperfusion

The first report of apoptotic cell death during hepatic ischemia/reperfusion appeared in 1996.<sup>24</sup> Similar reports of postischemic apoptosis have appeared for heart, brain, and other organs.<sup>23,62,63</sup> In liver after 60 minutes of warm ischemia, the number of apoptotic hepatocytes evaluated by nuclear morphology increases during the first 24 hours of reperfusion.<sup>24</sup> Using the TUNEL assay, another study identified apoptotic hepatocytes in human allografts after transplantation.<sup>64</sup> Subsequent studies reported that sinusoidal endothelial cells undergo apoptosis during cold ischemia/reperfusion<sup>25</sup> and that both sinusoidal endothelial cells and hepatocytes undergo apoptosis after warm ischemia/reperfusion.<sup>65</sup> These observations were based largely on fluorescence TUNEL assays and DNA laddering in gels. In addition, electron microscopy showed that single cells meet the morphological definition of apoptosis.<sup>25,65</sup> Further, pancaspase inhibitors attenuated reperfusion injury after warm and cold ischemia.<sup>66,67</sup> By the criteria of TUNEL labeling, 60% to 80% of sinusoidal endothelial cells and hepatocytes undergo apoptosis within 6 hours of reperfusion.<sup>65,67</sup> Further studies suggested that Kupffer cells and platelets are responsible for inducing apoptosis in these liver cells through the release of TNF- $\alpha$ .<sup>68,69</sup>

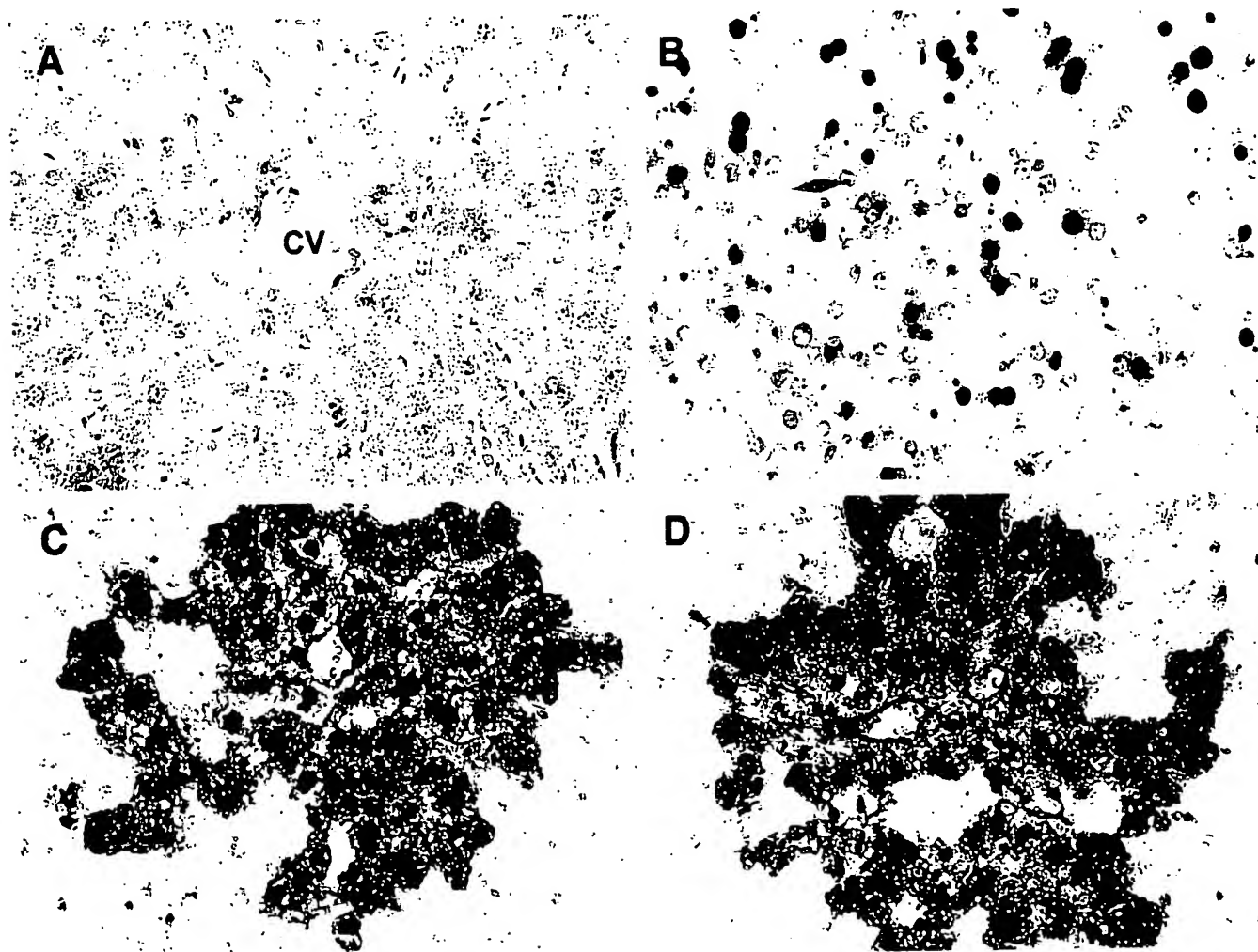
Despite the growing literature on apoptotic cell death after hepatic ischemia/reperfusion, concerns exist regarding the interpretation of these results and the relevance of apoptosis in the pathophysiology of reperfusion injury. The onset of necrotic cell death as judged by enzyme release and staining with trypan blue and propidium iodide occurs within minutes of reperfusion (see Figure 3).<sup>70-72</sup> After warm ischemia, reperfusion-induced oncotic necrosis occurs predominantly in hepatocytes and is accompanied by enzyme release.<sup>71,73</sup> After cold ischemia

**THIS PAGE BLANK (USPTO)**

during liver preservation for transplantation, necrotic death occurs nearly exclusively in sinusoidal endothelial cells and is accompanied by relatively little enzyme release because of the much smaller cytoplasmic mass of the endothelial cells.<sup>7,70</sup> The extent of this reperfusion-induced necrotic cell killing correlates well with graft failure after transplantation. When strict morphological criteria in combination with the TUNEL assay are used, apoptosis of endothelial cells and hepatocytes after 45 to 120 minutes of warm ischemia can be confirmed, but quantitatively apoptosis never exceeds 2% of the liver cells at risk.<sup>74</sup> Furthermore, necrotic cell death, identified by cell swelling, karyorrhexis, karyolysis, and vacuolization, accounts for more than 90% of all cell death, although many of the necrotic cells are TUNEL positive.<sup>74</sup> The relatively minor component of apoptotic cell

death after reperfusion is consistent with several other reports.<sup>24,75-77</sup> The limited amount of apoptotic cell death also correlates with limited or absent activation of caspases.<sup>74</sup> In contrast, during Fas- and TNF receptor-induced apoptosis *in vivo*, which affects approximately 15%–30% of hepatocytes, caspase 3 activity levels increase 10- to 20-fold or more, and extensive processing of procaspase 3 occurs<sup>78,79</sup>—features that are nearly absent after ischemia/reperfusion.<sup>74</sup>

Characteristically, apoptosis occurs in individual isolated cells. Even if large numbers of hepatocytes are induced *in vivo* to undergo apoptosis after activation of Fas or TNF receptors, individual cells rather than groups of contiguous cells show apoptotic features (Figure 4B). In contrast, oncotic necrosis typically occurs in confluent areas of adjacent cells (Figure 4C and D).<sup>61,80,81</sup> After



**Figure 4.** Hepatic TUNEL staining in TNF- $\alpha$ -induced apoptosis and acetaminophen-induced oncotic necrosis. Mice were untreated (A), received 700 mg/kg galactosamine and 100  $\mu$ g/kg endotoxin for 6 hours (B), or were treated with 300 mg/kg acetaminophen for 4 hours (C) or 6 hours (D). The assay shows a selective nuclear staining of individual hepatocytes during apoptosis (B) compared with the nuclear/cytosolic staining of contiguous cells in the centrilobular region of early oncotic necrosis after acetaminophen (C). Because of the more extensive karyolysis at later stages of oncotic necrosis, nuclear staining is less prominent in most cells compared with the cytosolic staining after 6 hours of acetaminophen treatment (D). CV, central vein.

**THIS PAGE BLANK (USPTO)**

hepatic ischemia/reperfusion, this necrosis typically occurs in the pericentral and midzonal regions of the hepatic lobule, because these regions are furthest removed from the oxygen supply.<sup>5,74</sup> These observations are consistent with the conclusion that the main mode of cell death during reperfusion injury is oncotic necrosis.

Another argument for apoptosis as the principal mode of cell killing after ischemia/reperfusion is the protective effect of Bcl-2 overexpression.<sup>82</sup> Bcl-2 interrupts apoptotic signaling at the level of the mitochondria<sup>38</sup> and prevents Fas-induced hepatocellular apoptosis.<sup>83</sup> However, Bcl-2 overexpression also inhibits necrotic cell death in hepatocytes and other cell types, possibly by inhibiting the MPT.<sup>84,85</sup> Oncosis can have other similarities to apoptosis, such as translocation of Bax to the mitochondria<sup>86</sup> and release of mitochondrial cytochrome *c* (without caspase 3 activation)<sup>87</sup> during acetaminophen-induced oncotic necrosis. Thus, identification of apoptosis as the principal mode of cell death requires evaluation of several parameters, which should qualitatively and quantitatively correlate with the extent of the assumed apoptosis.

Despite the predominance of necrosis over apoptosis after hepatic ischemia/reperfusion, several groups report protection by caspase inhibitors during ischemia/reperfusion.<sup>66,67</sup> However, protection may be rather modest, even with potent pancaspase inhibitors. For example, pancaspase inhibitors delay liver graft failure after prolonged cold ischemic storage by only approximately a day, without improvement of long-term graft survival.<sup>67</sup> This small and ultimately clinically irrelevant prolongation of survival may be due to anti-inflammatory effects, because pancaspase inhibitors block interleukin-1-converting enzyme (later renamed caspase 1), an enzyme involved in activating interleukin-1 and some other proinflammatory cytokines.<sup>88</sup> Apoptosis in a pathophysiological setting often promotes inflammation,<sup>89</sup> which in turn can extend and accelerate tissue injury.<sup>17</sup> During *Listeria* infection, hepatocellular apoptosis can promote neutrophil recruitment into the liver.<sup>90</sup> Moreover, TNF-induced parenchymal apoptosis triggers neutrophil transmigration and massive aggravation of the injury in an endotoxemia model.<sup>91,92</sup> Although the exact signaling mechanisms are not completely understood, apoptotic hepatocytes generate CXC chemokines,<sup>93</sup> which can signal neutrophil infiltration. A proinflammatory role of apoptosis in hepatic ischemia/reperfusion injury is also implied by findings that pancaspase inhibition decreases neutrophil recruitment into the liver, with attenuation of reperfusion injury.<sup>94</sup> Thus, apoptosis, even if limited to a relatively small number of cells, still has the potential to

affect overall injury by contributing to the amplification of the inflammatory response.

## Necrapoptosis

Part of the confusion concerning the roles of apoptosis and necrosis in ischemia/reperfusion and other forms of hepatic injury arises from the assumption that apoptotic and necrotic mechanisms are distinct and separate when, in fact, these mechanisms can be shared. In particular, the MPT plays an important role in oncotic necrosis, as well as in apoptosis. In ischemia, anaerobic glycolysis and ATP hydrolysis during ischemia rapidly decrease tissue pH, which protects strongly against necrotic cell killing despite profound ATP depletion.<sup>95,96</sup> Payback occurs when physiological pH returns after reperfusion, and the recovery of normal intracellular pH is an independent factor for precipitating lethal cellular reperfusion injury.<sup>70,71,97</sup> The mechanism of pH-dependent reperfusion injury involves onset of the MPT, a phenomenon that is inhibited by pH <7. Initially after reperfusion of hepatocytes in a cell culture model, mitochondria begin to repolarize, but as the intracellular pH approaches 7, mitochondria undergo inner membrane permeabilization, depolarization, and large-amplitude swelling.<sup>72</sup>

After onset of the MPT, mitochondrial uncoupling and activation of the mitochondrial uncoupler-stimulated adenosine triphosphatase lead to profound ATP depletion and ATP depletion-dependent necrotic cell death.<sup>72</sup> Cyclosporin A, a specific inhibitor of the MPT, prevents MPT-induced mitochondrial depolarization, inner membrane permeabilization, and ATP exhaustion after reperfusion and blocks the necrotic cell killing that ensues. The importance of ATP depletion is illustrated by the ability of the ATP-generating glycolytic substrate fructose to prevent this necrotic cell death. Only 15% to 20% of normal ATP is sufficient to prevent such necrotic cell killing.<sup>98,99</sup> Cytoprotection by fructose is downstream of the MPT, because fructose does not prevent the mitochondrial depolarization and inner membrane permeabilization induced by MPT-inducing treatments.<sup>100</sup> MPT-dependent necrotic cell death in models of ischemia/reperfusion to cultured hepatocytes is not blocked by caspase inhibitors and occurs without TUNEL staining.<sup>101</sup> The absence of TUNEL staining may reflect the release and dilution into the medium of nucleases after plasma membrane permeabilization.

However, when necrotic cell death is prevented by fructose, caspase 3-dependent apoptosis occurs instead, as documented by nuclear morphology, TUNEL labeling, and caspase activation.<sup>101</sup> Cyclosporin A, a specific

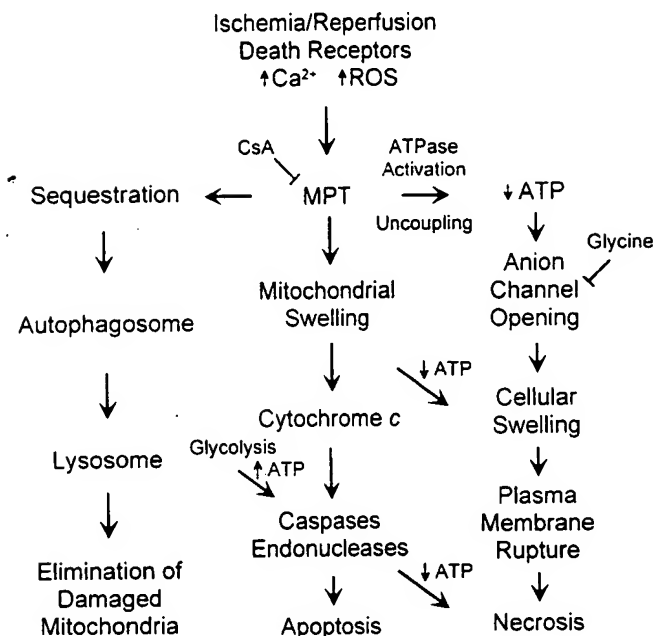
**THIS PAGE BLANK (USPTO)**

blocker of the MPT, still prevents this apoptosis, as do caspase inhibitors. Thus, the MPT is an obligatory event in both necrotic and apoptotic cell killing after ischemia/reperfusion. But how can one event, the MPT, lead to two such disparate events?

The onset of MPT leads to large-amplitude mitochondrial swelling, rupture of the outer membrane, and release of cytochrome *c* and other proteins from the intermembrane space between the mitochondrial inner and outer membranes. As discussed previously, cytochrome *c* interacts with apoptosis-activating factor-1 to promote caspase 9 activation, which then activates caspase 3. However, cytochrome *c*-dependent activation of caspase 9 requires ATP or the less abundant deoxyadenosine triphosphate.<sup>42</sup> Accordingly, the presence or absence of ATP can act as a "switch" between apoptosis and necrosis.<sup>102–104</sup> When reperfusion leads to both MPT onset and ATP depletion, apoptotic signaling is blocked at the level of the apoptosome, and necrosis occurs as a direct result of the failure of ATP regeneration (Figure 5). By contrast, if glycolytic substrate is available, profound ATP depletion is prevented, and necrosis does not occur. Instead, ATP-dependent apoptotic signaling occurs that is initiated by cytochrome *c* release after mitochondrial swelling. In cell-free extracts, the apparent  $K_M$  of ATP for activating apoptosomes is approximately 0.4 mmol/L,<sup>105,106</sup> which is only approximately 10% of the ATP concentration of normoxic hepatocytes. Thus, the amount of ATP needed to prevent necrosis (15%–20% of normoxic levels) is more than enough to permit cytochrome *c*-dependent caspase 9 and caspase 3 activation. Similarly, in hepatocytes exposed to calcium ionophore, the balance between ATP depletion after the MPT and ATP generation by glycolysis determines whether necrotic or apoptotic cell death occurs.<sup>100</sup> Thus, one mitochondrial event, the MPT, can lead to both apoptosis and necrosis (Figure 5). Consistent with these *in vitro* findings, the mode of hepatic cell death can be apoptosis after resuscitation after a shorter period of hemorrhagic shock when cellular ATP levels fully recover but can be necrotic after a longer period of shock when ATP levels remain suppressed after resuscitation.<sup>102</sup>

Just as a necrotic process can be converted to an apoptotic one, a process that starts with classic apoptotic signaling may switch to necrosis if ATP depletion or another change leads to breakdown of the plasma membrane permeability barrier. Apoptosis resulting in such secondary necrosis typically occurs during Fas antibody-induced hepatocellular injury *in vivo*.<sup>34</sup> In Fas receptor ligation-dependent liver injury, mitochondrial cytochrome *c* release, activation of the caspase cascade, DNA

## Necrapoptosis



**Figure 5.** Scheme of mitochondrial permeability transition (MPT)-dependent events in necrapoptosis. Ischemia/reperfusion, death-receptor activation, mitochondrial  $\text{Ca}^{2+}$  loading, and reactive oxygen species (ROS) are some of the events that promote onset of the MPT. If MPT onset occurs in relatively few mitochondria, the organelles become sequestered into autophagosomes for lysosomal digestion, a process that eliminates the damaged and potentially toxic mitochondria. When the MPT involves more mitochondria, mitochondrial swelling leads to outer membrane rupture and cytochrome *c* release. Provided that ATP is available from glycolysis and still-intact mitochondria, cytochrome *c* activates downstream caspases and other executioner enzymes of apoptosis. When MPT onset is abrupt and involves most mitochondria, ATP becomes profoundly depleted, which blocks caspase activation. Instead, ATP depletion leads to the opening of a glycine-sensitive organic anion channel to initiate a metastable state that culminates with plasma membrane rupture and the onset of necrotic cell death. If ATP depletion occurs during downstream apoptotic signaling, then secondary necrosis may supervene. CsA, cyclosporin A.

fragmentation, and morphological changes characteristic of apoptosis occur before the onset of secondary necrosis, with hepatocellular enzyme release and inflammatory changes.<sup>34,78,79</sup> Thus, in an apoptotic process that ultimately culminates in secondary necrosis, apoptotic mechanisms nonetheless remain clearly identifiable by morphological and biochemical parameters at earlier times.

The ability of a necrotic process to be converted to an apoptotic one and vice versa illustrates that apoptotic and necrotic cell death are not necessarily distinct and independent events. To the contrary, pathways leading to necrosis and apoptosis can be shared, a phenomenon called *necrapoptosis* or *aponecrosis*.<sup>107,108</sup> In necrapoptosis, events such as the MPT initiate a chain reaction that

**THIS PAGE BLANK (USPTO)**

culminates in either apoptosis or necrosis, depending on other variables, such as ATP supply (Figure 5). By inducing the MPT, ischemia/reperfusion causes both apoptosis and necrosis, although in a particular circumstance one or the other may predominate. The concept of necrapoptosis explains why features of apoptosis and necrosis often coexist in liver and other tissues, especially after pathologic insults such as ischemia/reperfusion or drug-induced liver injury. Recent work also suggests that limited onset of the MPT induces autophagy, a process by which effete, damaged, or superfluous mitochondria and other organelles are eliminated from cells by lysosomal degradation.<sup>109</sup> Thus, the concept of MPT-dependent necrapoptosis explains how injury can progress from reversible changes associated with tissue repair to apoptosis and then to necrosis. When the MPT occurs in only a few mitochondria, autophagy is stimulated, and the involved mitochondria are segregated for lysosomal degradation without stimulation of apoptotic signaling. With greater injury and more widespread MPT induction, apoptosis develops because of cytochrome *c*-dependent caspase activation. With even greater injury and MPT induction, ATP decreases to levels that no longer support apoptotic signaling, and oncotic necrosis develops instead (Figure 5).

## Conclusions

In liver, oncotic necrosis and apoptosis share features and mechanisms. DNA degradation after necrosis causes TUNEL labeling, which may be incorrectly interpreted as apoptotic cell death. During apoptosis in pathophysiological settings, inflammatory responses and enzyme release occur that resemble a necrotic process. Frequently, oncotic necrosis and apoptosis coexist after toxic, hypoxic, and inflammatory liver injury. The coexistence of the 2 patterns of cell death likely reflects shared mechanistic pathways. Experimental or clinical settings will determine whether cells die predominantly by apoptosis or oncotic necrosis. Therefore, it is important to evaluate critical cell death pathways under clinically relevant conditions to discover new therapeutic intervention strategies in hepatic ischemia/reperfusion injury.

## References

- Strasberg SM. Preservation injury and donor selection: it all starts here. *Liver Transpl Surg* 1997;5(suppl 1):S1-S7.
- Jaeschke H. Mechanisms of reperfusion injury after warm ischemia of the liver. *J Hepatobiliary Pancreat Surg* 1998;21:402-408.
- Bilzer M, Gerbes AL. Preservation injury of the liver: mechanisms and novel therapeutic strategies. *J Hepatol* 2000;32:508-515.
- Menger MD, Vollmar B. Role of microcirculation in transplantation. *Microcirculation* 2000;7:291-306.
- Lemasters JJ. Hypoxic, ischemic, and reperfusion injury to liver. In: Arias IM, ed. *The liver: biology and pathobiology*. 4th ed. Philadelphia: Lippincott Williams & Wilkins, 2001:257-279.
- Jaeschke H. Molecular mechanisms of hepatic ischemia-reperfusion injury and preconditioning. *Am J Physiol* 2003;284:G15-G26.
- Caldwell-Kenkel JC, Currin RT, Tanaka Y, Thurman RG, Lemasters JJ. Reperfusion injury to endothelial cells following cold storage of rat liver. *Hepatology* 1989;10:292-299.
- Lemasters JJ. The mitochondrial permeability transition and the calcium, oxygen and pH paradoxes: one paradox after another. *Cardiovasc Res* 1999;44:470-473.
- Lemasters JJ, Caldwell-Kenkel JC, Currin RT, Tanaka Y, Marzi I, Thurman RG. Endothelial cell killing and activation of Kupffer cells following reperfusion of rat livers stored in Euro-Collins solution. In: Wisse E, Knook DL, Decker K, eds. *Cells of the hepatic sinusoid*. Volume 2. Rijswijk, The Netherlands: Kupffer Cell Foundation, 1989:277-280.
- Jaeschke H, Farhood A. Neutrophil and Kupffer cell-induced oxidant stress and ischemia-reperfusion injury in rat liver *in vivo*. *Am J Physiol* 1991;260:G355-G362.
- Jaeschke H, Bautista AP, Spolarics Z, Spitzer JJ. Superoxide generation by Kupffer cells and priming of neutrophils during reperfusion after hepatic ischemia. *Free Radic Res Commun* 1991;15:277-284.
- Lindert KA, Caldwell-Kenkel JC, Nukina S, Lemasters JJ, Thurman RG. Activation of Kupffer cells on reperfusion following hypoxia: particle phagocytosis in a low-flow, reflow model. *Am J Physiol* 1992;262:G345-G350.
- Liu P, McGuire GM, Fisher MA, Farhood A, Smith CW, Jaeschke H. Activation of Kupffer cells and neutrophils for reactive oxygen formation is responsible for endotoxin-enhanced liver injury after hepatic ischemia. *Shock* 1995;3:56-62.
- Lichtman SN, Lemasters JJ. Role of cytokines and cytokine-producing cells in reperfusion injury to the liver. *Semin Liver Dis* 1999;19:171-187.
- Jaeschke H, Farhood A, Bautista AP, Spolarics Z, Spitzer JJ. Complement activates Kupffer cells and neutrophils during reperfusion after hepatic ischemia. *Am J Physiol* 1993;264:G801-G809.
- Colletti LM, Remick DG, Burtch GD, Kunkel SL, Strieter RM, Campbell DA Jr. Role of tumor necrosis factor- $\alpha$  in the pathophysiologic alterations after hepatic ischemia/reperfusion injury in the rat. *J Clin Invest* 1990;85:1936-1943.
- Jaeschke H, Farhood A, Smith CW. Neutrophils contribute to ischemia/reperfusion injury in rat liver *in vivo*. *FASEB J* 1990;4:3355-3359.
- Jaeschke H, Bautista AP, Spolarics Z, Spitzer JJ. Superoxide generation by neutrophils and Kupffer cells during *in vivo* reperfusion after hepatic ischemia in rats. *J Leukoc Biol* 1992;52:377-382.
- Jaeschke H, Farhood A, Bautista AP, Spolarics Z, Spitzer JJ, Smith CW. Functional inactivation of neutrophils with a Mac-1 (CD11b/CD18) monoclonal antibody protects against ischemia-reperfusion injury in rat liver. *Hepatology* 1993;17:915-923.
- Mavier P, Preaux A-M, Guigui B, Lescs MC, Zafrani ES, Dhumeaux D. *In vitro* toxicity of polymorphonuclear neutrophils to rat hepatocytes: evidence for a proteinase-mediated mechanism. *Hepatology* 1988;8:254-258.
- Pannen BH, Al-Adili F, Bauer M, Clemens MG, Geiger KK. Role of endothelins and nitric oxide in hepatic reperfusion injury in the rat. *Hepatology* 1998;27:755-764.
- Vollmar B, Glasz J, Leiderer R, Post S, Menger MD. Hepatic microcirculatory perfusion failure is a determinant of liver dysfunction in warm ischemia-reperfusion. *Am J Pathol* 1994;145:1421-1431.
- Gottlieb RA, Burleson KO, Kloner RA, Babior BM, Engler RL.

**THIS PAGE BLANK (USPTO)**

- Reperfusion injury induces apoptosis in rabbit cardiomyocytes. *J Clin Invest* 1994;94:1621-1628.
24. Sasaki H, Matsuno T, Tanaka N, Orita K. Activation of apoptosis during the reperfusion phase after rat liver ischemia. *Transplant Proc* 1996;28:1908-1909.
  25. Gao W, Bentley RC, Madden JF, Clavien PA. Apoptosis of sinusoidal endothelial cells is a critical mechanism of preservation injury in rat liver transplantation. *Hepatology* 1998;27:1652-1660.
  26. Trump BF, Goldblatt PJ, Stowell RE. Studies of necrosis in vitro of mouse hepatic parenchymal cells. Ultrastructural alterations in endoplasmic reticulum, Golgi apparatus, plasma membrane, and lipid droplets. *Lab Invest* 1965;14:2000-2028.
  27. Lemasters JJ, DiGuiseppi J, Nieminen A-L, Herman B. Blebbing, free  $Ca^{2+}$  and mitochondrial membrane potential preceding cell death in hepatocytes. *Nature* 1987;325:78-81.
  28. Nieminen AL, Gores GJ, Wray BE, Tanaka Y, Herman B, Lemasters JJ. Calcium dependence of bleb formation and cell death in hepatocytes. *Cell Calcium* 1988;9:237-246.
  29. Zahrebeliski G, Nieminen AL, Al-Ghoul K, Qian T, Herman B, Lemasters JJ. Progression of subcellular changes during chemical hypoxia to cultured rat hepatocytes: a laser scanning confocal microscopic study. *Hepatology* 1995;21:1361-1372.
  30. Nishimura Y, Lemasters JJ. Glycine blocks opening of a death channel in cultured hepatic sinusoidal endothelial cells during chemical hypoxia. *Cell Death Differ* 2001;8:850-858.
  31. Majno G, Joris I. Apoptosis, oncosis, and necrosis. An overview of cell death. *Am J Pathol* 1995;146:3-15.
  32. Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* 1972;26:239-257.
  33. Camilleri-Broet S, Vanderwerff H, Caldwell E, Hockenbery D. Distinct alterations in mitochondrial mass and function characterize different models of apoptosis. *Exp Cell Res* 1998;239:277-292.
  34. Ogasawara J, Watanabe-Fukunaga R, Adachi M, Matsuzawa A, Kasugai T, Kitamura Y, Itoh N, Suda T, Nagata S. Lethal effect of the anti-Fas antibody in mice. *Nature* 1993;364:806-809.
  35. Peter ME, Krammer PH. Mechanisms of CD95 (APO-1/Fas)-mediated apoptosis. *Curr Opin Immunol* 1998;10:545-551.
  36. Vieira HL, Kroemer G. Pathophysiology of mitochondrial cell death control. *Cell Mol Life Sci* 1999;56:971-976.
  37. Lemasters JJ, Qian T, He L, Kim J-S, Elmore SP, Cascio WE, Brenner DA. Role of mitochondrial inner membrane permeabilization in necrotic cell death, apoptosis and autophagy. *Antioxid Redox Signal* 2002;4:769-781.
  38. Gross A, McDonnell JM, Korsmeyer SJ. Bcl-2 family members and the mitochondria in apoptosis. *Genes Dev* 1999;13:1899-1911.
  39. Cohen GM. Caspases: the executioners of apoptosis. *Biochem J* 1997;326:1-16.
  40. Scaffidi C, Fulda S, Srinivasan A, Friesen C, Li F, Tomaselli KJ, Debatin KM, Krammer PH, Peter ME. Two CD95 (APO-1/Fas) signaling pathways. *EMBO J* 1998;17:1675-1687.
  41. Yin XM. Bid, a critical mediator for apoptosis induced by the activation of Fas/TNF-R1 death receptors in hepatocytes. *J Mol Med* 2000;78:203-211.
  42. Liu X, Kim CN, Yang J, Jemmerson R, Wang X. Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c. *Cell* 1996;86:147-157.
  43. Cande C, Cohen I, Daugas E, Ravagnan L, Larochette N, Zamzami N, Kroemer G. Apoptosis-inducing factor (AIF): a novel caspase-independent death effector released from mitochondria. *Biochimie* 2002;84:215-222.
  44. Du C, Fang M, Li Y, Li L, Wang X. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell* 2000;102:33-42.
  45. Verhagen AM, Ekert PG, Pakusch M, Silke J, Connolly LM, Reid GE, Moritz RL, Simpson RJ, Vaux DL. Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell* 2000;102:43-53.
  46. Bernardi P. Mitochondrial transport of cations: channels, exchangers, and permeability transition. *Physiol Rev* 1999;79:1127-1155.
  47. Shimizu S, Ide T, Yanagida T, Tsujimoto Y. Electrophysiological study of a novel large pore formed by Bax and the voltage-dependent anion channel that is permeable to cytochrome c. *J Biol Chem* 2000;275:12321-12325.
  48. Korsmeyer SJ, Wei MC, Saito M, Weiler S, Oh KJ, Schlesinger PH. Pro-apoptotic cascade activates BID, which oligomerizes BAK or BAX into pores that result in the release of cytochrome c. *Cell Death Differ* 2000;7:1166-1173.
  49. Wang X. The expanding role of mitochondria in apoptosis. *Genes Dev* 2001;15:2922-2933.
  50. Hatano E, Bradham CA, Stark A, Iimuro Y, Lemasters JJ, Brenner DA. The mitochondrial permeability transition augments Fas-induced apoptosis in mouse hepatocytes. *J Biol Chem* 2000;275:11814-11823.
  51. Bredesen DE. Apoptosis: overview and signal transduction pathways. *J Neurotrauma* 2000;17:801-810.
  52. Schwabe RF, Bennett BL, Manning AM, Brenner DA. Differential role of I kappa B kinase 1 and 2 in primary rat hepatocytes. *Hepatology* 2001;33:81-90.
  53. Chen G, Goeddel DV. TNF-R1 signaling: a beautiful pathway. *Science* 2002;296:1634-1635.
  54. Bradham CA, Qian T, Streetz K, Trautwein C, Brenner DA, Lemasters JJ. The mitochondrial permeability transition is required for tumor necrosis factor alpha-mediated apoptosis and cytochrome c release. *Mol Cell Biol* 1998;18:6353-6364.
  55. Bedner E, Li X, Gorczyca W, Melamed MR, Darzynkiewicz Z. Analysis of apoptosis by laser scanning cytometry. *Cytometry* 1999;35:181-195.
  56. Dong Z, Saikumar P, Weinberg JM, Venkatachalam MA. Internucleosomal DNA cleavage triggered by plasma membrane damage during necrotic cell death. Involvement of serine but not cysteine proteases. *Am J Pathol* 1997;151:1205-1213.
  57. Ray SD, Sorge CL, Raucy JL, Corcoran GB. Early loss of large genomic DNA in vivo with accumulation of  $Ca^{2+}$  in the nucleus during acetaminophen-induced liver injury. *Toxicol Appl Pharmacol* 1990;106:346-351.
  58. Shen W, Kamendulis LM, Ray SD, Corcoran GB. Acetaminophen-induced cytotoxicity in cultured mouse hepatocytes: effects of  $Ca(2+)$ -endonuclease, DNA repair, and glutathione depletion inhibitors on DNA fragmentation and cell death. *Toxicol Appl Pharmacol* 1992;112:32-40.
  59. Grasl-Kraupp B, Ruttkay-Nedecky B, Koudelka H, Bukowska K, Bursch W, Schulte-Hermann R. In situ detection of fragmented DNA (TUNEL assay) fails to discriminate among apoptosis, necrosis, and autolytic cell death: a cautionary note. *Hepatology* 1995;21:1465-1468.
  60. Ohno M, Takemura G, Ohno A, Misao J, Hayakawa Y, Minatoguchi S, Fujiwara T, Fujiwara H. "Apoptotic" myocytes in infarct area in rabbit hearts may be oncotic myocytes with DNA fragmentation: analysis by immunogold electron microscopy combined with in situ nick end-labeling. *Circulation* 1998;98:1422-1430.
  61. Gujral JS, Knight TR, Farhood A, Bajt ML, Jaeschke H. Mode of cell death after acetaminophen overdose in mice: apoptosis or oncotic necrosis? *Toxicol Sci* 2002;67:322-328.
  62. Banasiak KJ, Xia Y, Haddad GG. Mechanisms underlying hypoxia-induced neuronal apoptosis. *Prog Neurobiol* 2000;62:215-249.
  63. Sheridan AM, Bonventre JV. Cell biology and molecular mecha-

**THIS PAGE BLANK (USPTO)**

- nisms of injury in ischemic acute renal failure. *Curr Opin Nephrol Hypertens* 2000;9:427-434.
64. Borghi-Scoazec G, Scoazec JY, Durand F, Bernuau J, Belghiti J, Feldmann G, Henin D, DeGott C. Apoptosis after ischemia-reperfusion in human liver allografts. *Liver Transpl Surg* 1997;3:407-415.
  65. Kohli V, Seizner M, Madden JF, Bentley RC, Clavien PA. Endothelial cell and hepatocyte deaths occur by apoptosis after ischemia-reperfusion injury in the rat liver. *Transplantation* 1999;67:1099-1105.
  66. Cursio R, Gugenheim J, Ricci JE, Crenesse D, Rostagno P, Maulon L, Saint-Paul MC, Ferrua B, Auberger AP. A caspase inhibitor fully protects rats against lethal normothermic liver ischemia by inhibition of liver apoptosis. *FASEB J* 1999;13:253-261.
  67. Natori S, Seizner M, Valentino KL, Fritz LC, Srinivasan A, Clavien PA, Gores GJ. Apoptosis of sinusoidal endothelial cells occurs during liver preservation injury by a caspase-dependent mechanism. *Transplantation* 1999;68:89-96.
  68. Sindram D, Porte RJ, Hoffman MR, Bentley RC, Clavien PA. Platelets induce sinusoidal endothelial cell apoptosis upon reperfusion of the cold ischemic rat liver. *Gastroenterology* 2000;118:183-191.
  69. Sindram D, Porte RJ, Hoffman MR, Bentley RC, Clavien PA. Synergism between platelets and leukocytes in inducing endothelial cell apoptosis in the cold ischemic rat liver: a Kupffer cell-mediated injury. *FASEB J* 2001;15:1230-1232.
  70. Caldwell-Kenkel JC, Thurman RG, Lemasters JJ. Selective loss of nonparenchymal cell viability after cold, ischemic storage of rat livers. *Transplantation* 1988;45:834-837.
  71. Currin RT, Gores GJ, Thurman RG, Lemasters JJ. Protection by acidotic pH against anoxic cell killing in perfused rat liver: evidence for a 'pH paradox'. *FASEB J* 1991;5:207-210.
  72. Qian T, Nieminen AL, Herman B, Lemasters JJ. Mitochondrial permeability transition in pH-dependent reperfusion injury to rat hepatocytes. *Am J Physiol* 1997;273:C1783-C1792.
  73. Bradford BU, Marotto M, Lemasters JJ, Thurman RG. New simple models to evaluate zone-specific damage to hypoxia in the perfused rat liver: time course and effect of nutritional state. *J Pharmacol Exp Ther* 1986;236:263-268.
  74. Gujral JS, Bucci TJ, Farhood A, Jaeschke H. Mechanism of cell death during hepatic ischemia-reperfusion in rats: apoptosis or necrosis? *Hepatology* 2001;33:397-405.
  75. Redaelli CA, Tian YH, Schaffner T, Ledermann M, Baer HU, Dufour JF. Extended preservation of rat liver graft by induction of heme oxygenase-1. *Hepatology* 2002;35:1082-1092.
  76. Rentsch M, Beham A, Ilesalniks I, Mirwald T, Anthuber M, Jauch KW. Impact of prolonged cold ischemia and reperfusion on apoptosis, activation of caspase 3, and expression of bax after liver transplantation in the rat. *Transplant Proc* 2001;33:850-851.
  77. Ikebe N, Akaike T, Miyamoto Y, Hayashida K, Yoshitake J, Ogawa M, Maeda H. Protective effect of S-nitrosylated-protease inhibitor on hepatic ischemia-reperfusion injury. *J Pharmacol Exp Ther* 2000;295:904-911.
  78. Bajt ML, Lawson JA, Vonderfecht SL, Gujral JS, Jaeschke H. Protection against Fas receptor-mediated apoptosis in hepatocytes and nonparenchymal cells by a caspase-8 inhibitor *in vivo*: evidence for postmitochondrial processing of caspase-8. *Toxicol Sci* 2000;58:109-117.
  79. Bajt ML, Vonderfecht SL, Jaeschke H. Differential protection with inhibitors of caspase-8 and caspase-3 in murine models of tumor necrosis factor and Fas receptor-mediated hepatocellular apoptosis. *Toxicol Appl Pharmacol* 2001;175:243-252.
  80. Lawson JA, Fisher MA, Simmons CA, Farhood A, Jaeschke H. Inhibition of Fas receptor (CD95)-induced hepatic caspase activation and apoptosis by acetaminophen in mice. *Toxicol Appl Pharmacol* 1999;156:179-186.
  81. Gujral JS, Farhood A, Jaeschke H. Oncotic necrosis and caspase-dependent apoptosis during galactosamine-induced liver injury in rats. *Toxicol Appl Pharmacol* 2003;190:37-46.
  82. Bilbao G, Contreras JL, Eckhoff DE, Mikheeva G, Krasnykh V, Douglas JT, Thomas FT, Thomas JM, Curriel DT. Reduction of ischemia-reperfusion injury of the liver by *in vivo* adenovirus-mediated gene transfer of the antiapoptotic Bcl-2 gene. *Ann Surg* 1999;230:185-193.
  83. Lacroix V, Mignon A, Fabre M, Viollet B, Rouquet N, Molina T, Porteu A, Henrion A, Bouscary D, Varlet P, Joulin V, Kahn A. Bcl-2 protects from lethal hepatic apoptosis induced by an anti-Fas antibody in mice. *Nat Med* 1996;2:80-86.
  84. Yamabe K, Shimizu S, Kamiike W, Waguri S, Eguchi Y, Hasegawa J, Okuno S, Yoshioka Y, Ito T, Sawa Y, Uchiyama Y, Tsujimoto Y, Matsuda H. Prevention of hypoxic liver cell necrosis by *in vivo* human bcl-2 gene transfection. *Biochem Biophys Res Commun* 1998;243:217-223.
  85. Kane DJ, Ord T, Anton R, Bredesen DE. Expression of bcl-2 inhibits necrotic neural cell death. *J Neurosci Res* 1995;40:269-275.
  86. Adams ML, Pierce RH, Vail ME, White CC, Tonge RP, Kavanagh TJ, Fausto N, Nelson SD, Bruschi SA. Enhanced acetaminophen hepatotoxicity in transgenic mice overexpressing Bcl-2. *Mol Pharmacol* 2001;60:907-915.
  87. Knight TR, Jaeschke H. Acetaminophen-induced inhibition of Fas receptor-mediated liver cell apoptosis: mitochondrial dysfunction versus glutathione depletion. *Toxicol Appl Pharmacol* 2002;181:133-141.
  88. Siegmund B. Interleukin-1beta converting enzyme (caspase-1) in intestinal inflammation. *Biochem Pharmacol* 2002;64:1-8.
  89. Jaeschke H. Inflammation in response to hepatocellular apoptosis. *Hepatology* 2002;35:964-966.
  90. Rogers HW, Callery MP, Deck B, Unanue ER. *Listeria monocytogenes* induces apoptosis of infected hepatocytes. *J Immunol* 1996;156:679-684.
  91. Jaeschke H, Fisher MA, Lawson JA, Simmons CA, Farhood A, Jones DA. Activation of caspase-3 (CPP32)-like proteases is essential for TNF- $\alpha$ -induced hepatic parenchymal cell apoptosis and neutrophil-mediated necrosis in a murine endotoxin shock model. *J Immunol* 1998;160:3480-3486.
  92. Lawson JA, Fisher MA, Simmons CA, Farhood A, Jaeschke H. Parenchymal cell apoptosis as a signal for sinusoidal sequestration and transendothelial migration of neutrophils in murine models of endotoxin- and Fas-antibody-mediated liver injury. *Hepatology* 1998;28:761-767.
  93. Faouzi S, Burckhardt BE, Hanson JC, Campe CB, Schrum LW, Rippe RA, Maher JJ. Anti-Fas induces hepatic chemokines and promotes inflammation by an NF- $\kappa$ B-independent, caspase-3-dependent pathway. *J Biol Chem* 2001;276:49077-49082.
  94. Kobayashi A, Imamura H, Isobe M, Matsuyama Y, Soeda J, Matsunaga K, Kawasaki S. Mac-1 (CD11b/CD18) and intercellular adhesion molecule-1 in ischemia-reperfusion injury of rat liver. *Am J Physiol* 2001;281:G577-G585.
  95. Gores GJ, Nieminen AL, Fleishman KE, Dawson TL, Herman B, Lemasters JJ. Extracellular acidosis delays onset of cell death in ATP-depleted hepatocytes. *Am J Physiol* 1988;255:C315-C322.
  96. Gores GJ, Nieminen A-L, Wray BE, Herman B, Lemasters JJ. Intracellular pH during "chemical hypoxia" in cultured rat hepatocytes: protection by intracellular acidosis against the onset of cell death. *J Clin Invest* 1989;83:386-396.
  97. Bond JM, Chacon E, Herman B, Lemasters JJ. Intracellular pH and Ca<sup>2+</sup> homeostasis in the pH paradox of reperfusion injury to neonatal rat cardiac myocytes. *Am J Physiol* 1993;265:C129-C137.

THIS PAGE IS  
THIS PAGE BLANK (USPTO)

98. Anundi I, King J, Owen DA, Schneider H, Lemasters JJ, Thurman RG. Fructose prevents hypoxic cell death in liver. *Am J Physiol* 1987;253:G390-G396.
99. Nieminen AL, Saylor AK, Herman B, Lemasters JJ. ATP depletion rather than mitochondrial depolarization mediates hepatocyte killing after metabolic inhibition. *Am J Physiol* 1994;267:C67-C74.
100. Qian T, Herman B, Lemasters JJ. The mitochondrial permeability transition mediates both necrotic and apoptotic death of hepatocytes exposed to Br-A23187. *Toxicol Appl Pharmacol* 1999;154:117-125.
101. Kim J-S, Qian T, Lemasters JJ. ATP supply and the mitochondrial permeability transition in the switch from necrotic to apoptotic cell death after ischemia/reperfusion to cultured rat hepatocytes. *Gastroenterology* 2003;124:494-503.
102. Paxian M, Bauer I, Rensing H, Jaeschke H, Mautes AEM, Kolb SA, Wolf B, Stockhausen A, Jeblick S, Bauer M. Recovery of hepatocellular ATP and "pericentral apoptosis" after hemorrhage and resuscitation. *FASEB J* 2003;17:993-1002.
103. Richter C, Schweizer M, Cossarizza A, Franceschi C. Control of apoptosis by the cellular ATP level. *FEBS Lett* 1996;378:107-110.
104. Leist M, Single B, Castoldi AF, Kuhnle S, Nicotera P. Intracellular adenosine triphosphate (ATP) concentration: a switch in the decision between apoptosis and necrosis. *J Exp Med* 1997;185:1481-1486.
105. Eguchi Y, Shimizu S, Tsujimoto Y. Intracellular ATP levels determine cell death fate by apoptosis or necrosis. *Cancer Res* 1997;57:1835-1840.
106. Genini D, Budihardjo I, Plunkett W, Wang X, Carrera CJ, Cottam HB, Carson DA, Leoni LM. Nucleotide requirements for the in vitro activation of the apoptosis protein-activating factor-1-mediated caspase pathway. *J Biol Chem* 2000;275:29-34.
107. Lemasters JJ, V. Necrapoptosis and the mitochondrial permeability transition: shared pathways to necrosis and apoptosis. *Am J Physiol* 1999;276:G1-G6.
108. Formigli L, Papucci L, Tani A, Schiavone N, Tempestini A, Oriandini GE, Capaccioli S, Orlandini SZ. Aponecrosis: morphological and biochemical exploration of a syncratic process of cell death sharing apoptosis and necrosis. *J Cell Physiol* 2000;182:41-49.
109. Elmore SP, Qian T, Grissom SF, Lemasters JJ. The mitochondrial permeability transition initiates autophagy in rat hepatocytes. *FASEB J* 2001;15:2286-2287.

---

Received February 11, 2002. Accepted May 1, 2003.

Address requests for reprints to: Hartmut Jaeschke, Ph.D., Liver Research Institute, University of Arizona, College of Medicine, Room 6309, 1501 N. Campbell Avenue, Tucson, Arizona. e-mail: Jaeschke@email.arizona.edu; fax: (520) 626-5975.

Supported in part by National Institutes of Health grants ES06091 and AA12916 (to H.J.) and DK37034, DK59340, and AG13637 (to J.J.L.).

**THIS PAGE BLANK (USPTO)**



# ***STIC Search Report***

## ***Biotech-Chem Library***

**STIC Database Tracking Number: 106818**

**TO: Minh-Tam Davis  
Location: cm1/8a01/8e12  
Art Unit : 1642  
Tuesday, October 28, 2003**

**Case Serial Number: 10/016768**

**From : Susan Hanley  
Location: Biotech-Chem Library  
CM1 6B05  
Phone: 305-4053**

**susan.hanley@uspto.gov**

### **Search Notes**

**THIS PAGE BLANK (USPTO)**